Sarah Bell Blaise Gadanecz Leonardo Gambacorta Fernando Perez-Cruz Vatsala Shreeti

*Sarah.Bell@bis.org* *Blaise.Gadanecz@bis.org* *Leonardo.Gambacorta@bis.org* *Fernando.Perez-Cruz@bis.org* *Vatsala.Shreeti@bis.org*

# Online appendix to BIS Bulletin no 100: Artificial intelligence and human capital: challenges for central banks

## Appendix A: Some examples of LLM copilots used in central banks

Artificial intelligence (AI) and machine learning (ML) are likely to affect all aspects of central banking, albeit to different extents. According to a survey conducted by the Irving Fisher Committee (IFC) (Araujo et al (forthcoming)), over 150 use cases of AI are being explored at central banks, with a quarter of central banks having at least one use case in production. The Central Banking Technology Leadership Forum, held in December 2024, devoted two of its three sessions to AI to showcase current use cases. Many use cases focused on the utilisation of ML models for forecasting or nowcasting, large language model (LLM)-based chatbots to enhance human capabilities and the ability to query a specific document for information. Based on the presentations at the forum and at the 4th IFC and Bank of Italy Workshop on "Data science in central banking", we provide some examples of AI and ML applications that central banks are exploring.

**Microsoft 365 and GitHub copilots at the Bank of Thailand (BoT):** The BoT has conducted extensive analysis of the uses of Microsoft 365 Copilot for various applications, ranging from document summarisation to internal chatbots for BoT employees. The copilot performs well at summarising documents, saving time and reducing biases. In addition, the BoT has started using GitHub Copilot for coding, increasing the efficiency of the BoT's software developers and data scientists. Two major issues continue to limit copilot applications. First, copilots struggle with conversations that keep switching between languages (English and Thai, in this case). Second, they sometimes produce hallucinations or fail to accurately retrieve the relevant information. Because of these limitations, the BoT decided not to use them for meeting transcription and compliance assessment (Yampratoom (2024)).

**TIA (Textbasierte Intelligente Assistenten) platform at the Deutsche Bundesbank:** The Bundesbank utilises an in-house version of GPT-4o, distributed by Microsoft, to manage private and confidential information. The TIA platform does not function directly as a chatbot; instead, it is enhanced with a template interface to support repeated queries, thereby facilitating the automation of tasks. The template is used to craft the prompt before querying the GPT-4o engine and includes several fields similar to an online form. It also allows for documents to be uploaded in the prompt. The documents are indexed using a retrieval-augmented generation (RAG) system[1] to enable specific queries related to the uploaded document. The TIA platform encourages users to share their templates with other Bundesbank employees, increasing the replicability of the use cases in the platform and sharing working prompts across the organisation (Blankenburg and Röhe (2024)).

**Retrieval-augmented generation (RAG) copilot at the Federal Reserve Board (FRB):** The FRB, conducted a systematic experiment to validate a RAG copilot solution for answering specific questions based on commercial banks' financial documents. The results indicate that the time required to gather information decreased from five to six minutes when a human searches for the answer in the documents to 30–40 seconds per query on the RAG copilot, with a 20% increase in the quality of the answers. Furthermore, annotators with experience using AI, but without expertise in a particular central banking

[1] Retrieval-augmented generation (RAG) is an AI model that combines retrieval-based and generative approaches. It retrieves relevant information from a large data set and then generates coherent, contextually accurate responses, enhancing the quality and accuracy of generated text.

area, were able to outperform experts who were not familiar with AI tools. This underscores the importance of training central bank staff in the use of these tools. At the same time, the increase in quality was observed only for simple questions, and the RAG copilot solution could not provide better answers to more complex questions. This points to the need for human supervision while using such copilots and the need for the tools themselves to be improved (Botti et al (2025)).

**Processing survey data at the Reserve Bank of Australia (RBA):** The RBA has identified that AI can assist with processing the Bank's surveys. The RBA tested a new process in which a data scientist and a survey expert collaborate to maximise the value derived from each survey. The process is incremental: it allows for both obtaining a general overview of the survey quickly and allowing for further enhancement of the ML algorithms and human involvement if more detailed information is required. This experience demonstrates that cooperation between data scientists and domain experts can enhance the insights gained from AI use (Gray and Jones (2025)).

**Deal backing check at the BIS**: The Banking department at the BIS has developed a back-office deal backing check AI solution to automatically verify tens of thousands of yearly transactions, ensuring that trade details agreed to between parties (eg instrument, price, quantity, settlement instructions) match the data entered into IT systems. This LLM application reconciles unstructured conversations between BIS traders and counterparty traders, which occur over Bloomberg, email, Refinitiv etc, with the data captured in trading systems. Implementing this AI solution can save significant time for back-office employees when validating trades for various types of financial products, such as foreign exchange (FX) spots, FX derivatives, deposits etc. Currently plans are to use the AI application in copilot mode, assisting back-office employees. Eventually, depending on the performance of the model and the implementation of strong complementary controls, it might be used in an autonomous mode.

## Appendix B: New roles in central banks for AI integration and changes in existing roles

New staff roles will become increasingly important in central banks; we outline some of them below. Some of these roles may already exist in central banks, but their importance will increase as AI is adopted.

**Machine learning (ML) researcher**. ML researchers will customise foundational multimodal models to meet the central bank's needs. These individuals will possess expertise in efficiently adapting general models for specific applications. They will determine how to train the models (eg in-context learning, fine-tuning or reinforcement learning with human feedback), what data to gather and how to align the model with the specific use (eg selecting the combination of valid loss functions).

**ML operations (ML ops) engineer**. ML ops engineers are responsible for deploying ML or generative AI (gen AI) models to ensure seamless operation. They will monitor the models' performance and retrain them according to the guidelines provided by the ML researchers. Gen AI models are complex systems that require effective integration between data and computational resources. If a copilot or agent takes too long to respond, it would not be usable; ensuring prompt responsiveness will be ML ops engineers' responsibility.

**Data engineer.** Data engineers will be responsible for maintaining data in a machine-readable format for gen AI and agents to access. Ensuring that the data are accompanied by the correct metadata and stored systematically will be paramount for training and utilising these models. Data engineers will generally be responsible for data integration and quality, particularly in handling unstructured data and aligning them with structured data.

**AI ethics officer.** AI ethics officers will ensure that the deployment and use of AI systems adhere to ethical guidelines and regulatory standards. They will be responsible for monitoring AI applications for biases, ensuring transparency and addressing any ethical concerns that arise. This role is crucial in maintaining public trust and aligning AI use with the central bank's values.

**AI trainer.** AI trainers will be responsible for educating and training central bank employees to effectively use AI tools and systems. They will develop training programs, create instructional materials and provide hands-on training sessions. Their role will be essential in both scenarios to ensure that staff can use AI technologies effectively.

At the same time, the use of AI copilots will also affect existing roles:

- **Analysts and statisticians:** Traditionally, analysts spend significant time collecting and processing data, running models and generating reports. With LLM-based copilots, analysts can automate data collection and initial analysis, as well as get coding support, allowing them to focus more on interpreting results, conducting deeper analyses and developing strategic insights.
- **Financial supervisors:** Many central banks are already using supervisory technology (suptech) tools for financial risk assessment, data visualisation and automation of supervisory processes (Prenio (2024)). Such tools are being used to help supervisors identify credit exposures of small banks, to compare indicators across banks and to leverage granular data from credit registries. Supervisory copilots can enhance the efficiency of such tools by further automating data collection, processing vast amounts of data quickly and accurately, and identifying patterns and anomalies to detect risks and non-compliance in real time.
- **Human resource (HR) professionals:** Central bank HR departments can use AI to streamline administrative tasks and document generation and improve drafting of job profiles by analysing market trends and skill requirements. They can use AI chatbots for HR service delivery such as answering employee queries and scheduling interviews.

## Appendix C: Data, model and IT choices and workforce composition

**A central bank's workforce composition will be shaped by decisions on which gen AI models to use and how to collect data.** Preference for externally developed pre-trained models may be driven by expectations of lower development and maintenance costs; however, cost implications are difficult to predict. Central banks may also supplement internally generated or collected data with public or externally sourced data. Central banks that exclusively use externally procured models or data may need a smaller proportion of their workforce to have roles with a significant AI/ML-related component.

**The impact on workforce composition will also depend on the infrastructure choices of central banks.** Central banks could decide to (a) run proprietary systems on their internal servers, (b) use third-party tools in a cloud computing environment or (c) adopt hybrid approaches. These choices are less dependent on the scenarios themselves and more on the specific use cases of the copilots and agents. The decision to use cloud versus proprietary systems hinges primarily on data sensitivity and privacy concerns. For example, a central bank might use its proprietary systems on internal servers for its most sensitive data and applications, such as regulatory compliance monitoring, while employing cloud-based third-party tools for less sensitive tasks, like public sentiment analysis. These decisions would have an impact on workforce composition, particularly on the level of new technical expertise required and the extent to which existing staff need to be retrained or upskilled.

- **Proprietary systems on internal servers:** Central banks choosing to run proprietary AI systems on their internal servers would need to invest in significant in-house technical expertise and computational resources. This approach can provide greater control over data security and system customisation but requires robust IT infrastructure and skilled staff to manage the AI systems. There is an additional risk of "lock-in" to systems that become obsolete over time, leading to a trade-off between nimbleness and security. In scenario 1, internally developed copilots based on open weight models could utilise proprietary systems, ensuring that confidential data remain secure and are kept off public networks. In scenario 2, the proprietary systems could become more prevalent if agents are developed internally to manage confidential data and mitigate risks.

- **Third-party tools in a cloud computing environment:** Using open source or third-party AI tools hosted in a cloud computing environment can offer scalability and may reduce the need for extensive in-house IT infrastructure and specialised expertise. However, this approach may raise concerns about data privacy and security, as sensitive information is stored and processed off site. For example, LLMs can memorise parts of their training data and reproduce it when given specific prompts, compromising privacy (Carlini et al (2023)). Even with this approach, staff will need to understand the capabilities and limitations of third-party tools, as well as the operational risk that comes with their use. In scenario 1, copilots based on third-party tools (eg GitHub Copilot for software development) could be operated more efficiently in cloud environments, allowing access to the latest models. In scenario 2, cloud computing environments would be the preferred choice for AI agents that require internet access, especially if they do not handle confidential data or critical tasks.
- **Hybrid approaches:** Adopting a hybrid approach, where some AI systems are run on internal servers while others are hosted in the cloud, can provide a balance between control and flexibility.[2] An approach that combines the use of cloud and graphic processing units (GPUs), which are essential for processing large data sets and running complex AI models, may allow central banks to leverage the strengths of both proprietary and third-party solutions.[3] At the same time, staff would need to possess a diverse range of skills, including expertise in both on-premises and cloud-based technologies. Staff would also need to manage complex data workflows and make informed decisions about when to use proprietary systems versus third-party solutions.

## References


Araujo, D, R Schmidt, O Sirello, B Tissot and R Villarreal (forthcoming): "IFC report on governance and implementation of artificial intelligence in central banks: results of an IFC survey of central banks", *IFC Working Papers*.

Blankenburg, J and O Röhe (2024): "TIA platform: analysis of implementation of structural reforms", presentation at the Central Banking Technology Leadership Forum, 4–5 December.

Botti, N, F Haberkorn and S Khan (2025): "Efficacy of AI RAG tools for complex information extraction and data annotation tasks on bank documents", 4th IFC and Bank of Italy Workshop on Data Science in Central Banking, Rome, 18–20 February.

Carlini, N, D Ippolito, M Jagielski, K Lee, F Tramèr and C Zhang (2023): "Quantifying memorization across neural language models", International Conference on Learning Representations, 22 February.

Gray, N and D Jones (2025): "Transforming survey analysis: tools for central banks", 4th IFC and Bank of Italy Workshop on Data Science in Central Banking, Rome, 18–20 February.

Prenio, J (2024): "Peering through the hype – assessing suptech tools' transition from experimentation to supervision", *FSI Insights*, no 58, June.

Yampratoom, E (2024): "Experiences from implementing business use cases with generative AI", presentation at the Central Banking Technology Leadership Forum, 4–5 December.


[2] Cost considerations may also be a factor, though anecdotally, central banks have highlighted that the cost implications of various options remain unclear at this stage.

[3] A GPU is a specialised electronic circuit designed to rapidly process and render images, animations and videos. It is commonly used in computers and gaming consoles. GPUs also accelerate AI processing thanks to their parallel processing capabilities.