

Workshop 4

AI as a (financial) system shaper; implications for banks and supervision

Chair: Therese McCarthy Hockey, Deputy Chair, Australian Prudential Regulation Authority

In 2024, the Basel Committee published *Digitalisation of finance*, which described several ways in which banks were using or experimenting with artificial intelligence (AI) tools. Since then, a spate of advancements in AI technology has led to increasingly powerful AI systems. The capabilities of these AI systems continue to expand, and, by some metrics, to do so at an increasing rate. Banks, bank supervisors and other participants in and outside the financial sector are increasingly implementing and experimenting with AI tools. While the ultimate trajectory of AI's development is uncertain, it has the potential to dramatically alter the banking system and the broader economy. Supervisors and other authorities around the world are grappling with the promise and risk of the rapidly developing technology.

This workshop will focus on the ways in which AI technologies might shape the future of the banking sector, emphasising supervisory considerations related to banks' use of AI tools and AI-driven changes in the external environment.

AI tools present opportunities and risks for banks and supervisors

AI shows the potential to positively transform banks' businesses

Digitalisation of finance described the potential of AI tools to improve client experience, increase cost efficiencies and provide advantages in data processing.¹ At that time, many banks were cautiously experimenting with new AI tools in a variety of contexts. Since then, AI technology has developed rapidly, now surpassing human-level performance on some specialised benchmarks.² By some metrics, the capabilities of AI systems are doubling every few months and may be improving at an increasing rate.³ Increasingly, banks are using AI tools in their businesses, including implementation of maturing generative AI tools and increasingly piloting advanced agentic AI applications.

Banks are using or assessing the use of AI tools across front-, middle- and back-office applications. Some supervisors report that banks' use of AI is maturing, with AI tools being deployed in more material applications that are decision-critical and customer-facing.⁴ Uses include, but are not limited to:

¹ Basel Committee on Banking Supervision (2024). The report notes that AI tools can provide superior pattern recognition and predictive power, greater accuracy and consistency and greater ability to accommodate less-structured data.

² AI Security Institute (2026).

³ Kwa, T, et al.

⁴ McCarthy Hockey, T (2026).

- Know-your-customer and anti-money laundering compliance;
- Transaction monitoring and fraud detection;
- Credit scoring and underwriting;
- Software and systems modernisation and development, including in cybersecurity applications;
- Trading; and
- Direct customer interaction and engagement.

Some banks have begun to realise benefits from their use of AI. A recent report from the Cambridge Centre for Alternative Finance states that 40% of banks in their global sample have seen positive profitability impacts from the use of AI, while nearly all have seen positive productivity impacts.⁵ Various publicly available case studies provide more detail. For example, the FSB's consultation report, *Sounds Practices for Responsible Adoption of Artificial Intelligence (AI)*, describes one bank which embedded machine learning tools into its credit risk operations to identify early signs of increasing cashflow deterioration at certain borrowers. This enabled the bank to identify non-performing loans in a more accurate and timely manner.⁶ However, AI tools do not always bring the hoped-for benefits; the Cambridge report describes some institutions that have seen fewer or no improvements in their business via the use of AI tools.⁷

As AI technology continues to develop and access continues to broaden, many observers expect these use cases to multiply. Since the 2024 publication of *Digitalisation of finance*, banks have explored and some are implementing agentic AI tools. These AI tools can autonomously perform complex and multi-step tasks, working together and taking actions independently. Frontier AI models expand these capabilities even further, increasing the speed and scale with which tasks can be accomplished, demonstrating the ability to process and synthesise diverse types and sources of data and showing higher-level analytical and coding capabilities.

Banks' use of AI can pose prudential and financial stability risks

As banks adopt and experiment with the ways AI tools can enhance their businesses, they may be exposed to new and transformed risks. A variety of reports from supervisors and international organisations highlight these risks, some of which are discussed below.⁸

- **Data risks** related to model risk, bias, information security and performance. Weaknesses in the quality, integrity or suitability of data, as well as deficiencies in model design or calibration, could result in inaccurate, biased or unreliable outcomes.
- **Third party and concentration risks**, which can arise from banks' reliance on third-party service providers (TPSPs), such as providers of AI tools, in the performance of their critical operations.⁹ The use of TPSPs may reduce banks' direct control over their activities and assets, raising risks to banks' ability to conduct due diligence, monitor performance and manage data use. Concentration among TPSPs can increase financial stability risks (eg due to TPSP outage, herding behaviour caused by widespread use of the same models). These risks can arise when multiple services are provided by the same TPSP or a single or small number of TPSPs is providing services

⁵ Cambridge Centre for Alternative Finance (2026).

⁶ Financial Stability Board (2026b).

⁷ Cambridge Centre for Alternative Finance (2026).

⁸ This discussion of risks is drawn from several sources: Basel Committee on Banking Supervision (2024), Basel Committee on Banking Supervision (2025), Financial Stability Board (2024) and Financial Stability Board (2026).

⁹ See the definition of critical operations in Basel Committee on Banking Supervision (2021).

to a large number of banks. In the case of AI tools, the global reliance of banks on a few large AI companies could pose systemic and cross-border vulnerabilities.

- **Cyber risk and ICT risks** relating to the use of AI. For example, AI tools can be vulnerable to performance degradation from data poisoning attacks or to security breaches from prompt injections.¹⁰
- **Explainability challenges.** Explainability exists along a continuum and can be defined as the extent to which humans understand how an AI model or system generates outputs from inputs. Among other challenges, limited explainability can present problems in assessing the conceptual soundness of AI outputs and the compliance of AI tools.

These risks can be amplified by improper governance and human oversight procedures. Human oversight could be weakened by an overreliance on AI or an incomplete understanding of the ways in which AI tools operate. The use of “shadow AI,” or unauthorised use of AI tools, could exacerbate these risks and complicate risk management. All of these risks and challenges could be amplified by increasingly capable frontier AI models or models with a higher degree of autonomy, such as AI agents. As AI systems become more autonomous, there is also a risk that they optimise for assigned goals in unexpected ways, leading to unintended consequences or control failures.

The use of AI tools outside the banking sector can drive changes in the external environment

The use of AI tools outside the banking sector can also lead to changes in the risks that banks face.

One prominent example is the risk of **cyberattacks** facilitated by frontier AI models. As noted by governments and international organisations around the world, the current generation of frontier AI models have shown exceptional facility in bypassing cybersecurity defences and exploiting previously unknown vulnerabilities.¹¹ Many observers expect these capabilities to continue to develop.¹² In fact, in the months following the first demonstrations of frontier AI models, additional models have shown similar capabilities and agents from leading frontier AI labs have demonstrated the ability to successfully hack other organisations and work together in an attempt to deceive humans.¹³ These models can pose threats not only to banks, but to their clients and to other organisations including governments and infrastructure providers.

Similarly, the use of AI to facilitate **digital fraud and scams** has enhanced the capabilities and reduced the barriers to entry for fraudsters. Some supervisors report that the rise of AI tools has made it more challenging to detect frauds and easier for criminals to scale them.¹⁴ While estimates of the impact of digital fraud vary widely across jurisdictions, some estimate that global fraud losses exceeded USD 1

¹⁰ Data poisoning attacks are attacks in which incorrect or biased data are injected into models. Prompt injection involves the manipulation of AI tools using prompts to bypass safeguards and extract sensitive information.

¹¹ See, eg Five Eyes Cyber Security Agencies (2026).

¹² See, eg Bank of England, et al (2026).

¹³ In July 2026, an AI agent developed by OpenAI hacked into Hugging Face, an open-source AI development platform. See Hugging Face (2026). In August 2026, the AI Security Institute published a report detailing a series of unauthorised actions on the Internet by multiple independent AI agents coordinating to target real people and organisations; see AI Security Institute (2026).

¹⁴ See, eg Hong Kong Monetary Authority (2026b).

trillion in 2024 and that they are rapidly growing.¹⁵ As noted in a 2023 BCBS discussion paper, digital fraud can lead to costs and losses for banks.¹⁶

The impacts of rapid AI development and adoption in the external environment also pose direct **financial risks**. Industry participants are monitoring the large amount of debt financing used for AI-related capital expenditures and the potential for broader corrections in risk asset valuations driven by concerns related to AI valuations.¹⁷ At the same time, concern is mounting over exposures to industries – such as the software sector – which may be especially vulnerable to disruption from advances in AI. There are also concerns about financial system interconnections related to AI: private credit funds, which are interconnected with banks, insurers and other financial system participants, are heavily exposed to the software sector and could become an important source of AI-related financing.¹⁸

More broadly, the diffusion of AI across sectors has the potential to trigger a macroeconomic transition with uncertain consequences for labour markets. These impacts could vary across sectors and jurisdictions, creating unpredictable impacts. Under some scenarios, AI use could create additional uncertainty related to inflation levels and monetary policy transmission.¹⁹ If realised, these scenarios could drive further changes in the risk environment for banks.

Taken together, AI-driven shifts in the external environment may change the distribution of risks faced by banks across risk stripes, driving additional uncertainty and requiring adaptation from both banks and supervisors.

Supervisors are working with banks to weigh the benefits and risks of AI tools

AI tools have the potential to help banks deliver improved services in a more efficient manner and to manage changes in the external environment. For example, AI tools can be critical supports in helping banks and other organisations enhance their cyber defence capabilities.²⁰ With this in mind, many supervisors are actively encouraging banks in their jurisdictions to adopt AI to enhance and improve their businesses. Some supervisors have created AI “sandboxes” to facilitate AI experimentation and many supervisors and international organisations are conducting outreach to facilitate coordination among banks, supervisors and other industry participants.²¹ These and other actions reflect a view among both banks and supervisors that a failure to adapt to an AI-enabled world is a risk unto itself.

However, some supervisors have expressed concerns that governance and risk management processes are not keeping pace while banks rush to adopt AI technology.²² Concerns include concentration risks, a lack of AI expertise at the board level, and a lack of transparency over the use of AI within organisations.²³ At the same time, the rapidly changing risk environment requires a rapid response. To balance the need for a rapid response with the potential for risks associated with banks’ use of new AI

¹⁵ Khiaonarong, T and S Zheng (2026).

¹⁶ Basel Committee on Banking Supervision (2023).

¹⁷ Board of Governors of the Federal Reserve System (2026).

¹⁸ Cera, K, et al (2026).

¹⁹ Barhoumi, K, et al (2026).

²⁰ World Economic Forum (2026).

²¹ See, eg Hong Kong Monetary Authority (2026a).

²² McCarthy Hockey, T (2026).

²³ Australian Prudential Regulation Authority (2026).

tools, supervisors are working with banks to adopt AI in a responsible manner, focusing on proportional responses that emphasise the materiality of the use case within a given scenario.²⁴

Supervisors are also incorporating AI into their supervisory and monitoring processes. Authorities in several jurisdictions are using AI tools to enhance market surveillance, identify emerging risks and target supervisory efforts. Many supervisors are also adopting or planning to adopt AI tools for cyber security.²⁵

The development of AI technology may have dramatic and global implications for the banking system. The cross-border reach and multifarious impact of AI technologies highlight the importance of continued collaboration between banks, bank supervisors and other participants in and outside the financial sector as the banking system enters an AI-enabled future.

- Q1. How are the banks you supervise incorporating AI tools into their businesses? Are there notable differences in AI adoption between different types of banks (eg smaller banks versus large ones) or across jurisdictions?
- Q2. What are your supervisory expectations related to the use of AI, and how are these expectations evolving as AI capabilities and use cases mature?
- Q3. How is AI driving changes in the external environment? Do these changes differ across jurisdictions? How are banks and supervisors preparing for these changes?
- Q4. How are banks and supervisors striking a balance between rapid adoption of AI and ensuring that governance and risk management policies and processes are mature? How are you and the banks you supervise approaching proportionality and materiality in AI use?
- Q5. How are supervisors taking advantage of AI for supervision or financial stability monitoring?
- Q6. What is the outlook for the future rate of change? Are we entering an era in which the speed and level of change and uncertainty will be elevated?

References

Adrian, T (2026): "How Central Banks Can Contain Financial Stability Risks as AI Accelerates Change", July.

AI Security Institute (2026), "Incident Report: unsanctioned agent behaviour during cyber testing", August.

Australian Prudential Regulation Authority (2026), "APRA calls for a step-change in AI-related risk management and governance", April.

Bank of England, Financial Conduct Authority and HM Treasury (2026): "The Bank, FCA and HM Treasury joint statement on frontier AI models and cyber resilience", May.

Barhoumi, K, F de Carvalho, M Gorbanyov, Y Kido, D Koll, A D M Nguyen, D Ostojic, B Shang, N Tamirisa, S Toms and Y Zhao (2026): "Global Economic and Financial Implications of Artificial Intelligence: Lessons from a Scenario Planning Exercise; IMF Note No. 2026/002".

Basel Committee on Banking Supervision (2021): "Principles for operational resilience", March.

——— (2023): "Digital fraud and banking: supervisory and financial stability implications", November.

——— (2024): "Digitalisation of finance", May.

——— (2025), "Principles for the sound management of third-party risk", December.

²⁴ Bowman, M (2026).

²⁵ Adrian, T (2026).

Board of Governors of the Federal Reserve System (2026): *Financial Stability Report*, May.

Bowman, M (2026) "Artificial intelligence in the financial system", May.

Cambridge Centre for Alternative Finance (2026), *The 2026 Global AI in Financial Services Report*, (April).

Cera, K, D Dieckelmann, K Nikolov, G Schepens and O Blicke (2026) "Stress in global private credit markets and its implications for euro area financial stability", May.

Crowe, S, R Gupta, C Robinson, R Horne, D Imbordino, and N Andersen (2026): "Five Eyes cyber security agencies statement", June.

Financial Stability Board (2024): "The Financial Stability Implications of Artificial Intelligence", November.

——— (2026a):, Report on Vulnerabilities in Private Credit", May.

——— (2026b): "Sound Practices for Responsible Adoption of Artificial Intelligence Consultation Report", (June).

Hong Kong Monetary Authority (2026a): "Regulators launch GenA.I. Sandbox++ to foster A.I. innovation across financial services", March.

——— (2026b): *Supporting A.I. Adoption in Fighting Financial Crime*, June.

Hugging Face (2026): "Anatomy of a Frontier Lab Agent Intrusion: A Technical Timeline of the July 2026 Incident", July.

Khiaonarong, T and S Zheng (2026) "IMF Working Papers Volume 2026 Issue 062: The Rise of Cyber Events and Digital Fraud in the Financial Sector", March.

Kwa, T, B West, J Becker, A Deng, K Garcia, M Hasin, S Jawhar, M Kinniment, N Rush, S Von Arx, R Bloom, T Broadley, H Du, B Goodrich, N Jurkovic, L Miles, S Nix, T Lin, C Painter, N Parikh, D Rein, L J K Sato, H Wijk, D Ziegler, E Barnes and L Chan (Preprint): "Measuring AI Ability to Complete Long Software Tasks".

McCarthy Hockey, T (2026): "Fighting fire with fire", June.

Sajadieh, S, L Fattorini, R Perrault, Y Gil, V Parli, L Santarlaschi, J Pava, N Maslej, R Altman, E Brynjolfsson, C Brodley, J Clark, V Dignum, V Kumar, J Landay, T Lyons, J Manyika, J Niebles, Y Shoham, E Tabassi, R Wald, T Walsh and D Weld (2026): *The AI Index 2026 Annual Report*, April.

World Economic Forum (2026), *Empowering Defenders: AI for Cybersecurity*, May.