

Agents of change – speech by Sarah Breeden

Given at ECB Sintra Forum

Sarah Breeden sets out how AI is reshaping finance at speed. She explores how agentic AI is transforming cyber risk, markets and payments – and why central banks must adapt fast, strengthen resilience and cooperate globally to ensure the next technology surprise does not become a test of financial stability.

Speech

It's a privilege to be here to discuss AI and financial stability. It's hard to imagine a more timely or important issue. AI is reshaping finance at speed. And it's on us to ensure that the next technology surprise does not become a test of financial stability.

For some years, central banks, including the Bank of England, have worked to promote responsible AI adoption in the financial system.^[1] And since AI looks likely to be the next 'general purpose technology' and a key source of long-term economic growth^[2], it's in all our interests to do so.

But the apparent acceleration in AI capabilities, even compared with six months ago, creates a double challenge: not only enabling responsible adoption and managing risks, but also recognising that AI should transform how central banks do that job. We need to think as hard about the latter as the former.

How the technology is evolving

In 2019, the length of software task the latest AI models could complete doubled roughly every seven months.^[3] In 2024, that had sped up to around four months.^[4] Major advances this Spring in models for identifying cyber vulnerabilities suggest it may now be faster still.^[5]

So an already exponential increase in capability appears to be accelerating.^[6] We were surprised this Spring, and we should be prepared for further technology surprises.

These advances also suggest an inflection point. Around the start of this decade, generative AI systems created content on request. From late 2024, they were explicitly trained to reason through requests.^[7] Now with agentic AI, systems can autonomously chain together sequences of actions.

The financial system looks likely therefore to evolve quickly into one that operates more autonomously, at scale and speed: AI agents transacting on behalf of consumers and merchants; devising and executing trading strategies in financial markets; and identifying and chaining together cyber vulnerabilities.

Drivers of the transition and its financial stability risks

That will of course require central banks to respond. But before turning to that, I wanted to highlight that this transition is already macro-relevant, inherently uncertain and carries risks of its own. Four interdependent channels will shape its nature, scale and timing.

The first is AI's capability: for example, whether progress in recursive self-learning allows models to improve autonomously.

Second, capacity: whether compute keeps pace with capability and demand.

Third, diffusion: where adoption takes hold across sectors and tasks.

Fourth, how debt and equity finance fund this historically unprecedented pace of investment.

The Bank's Financial Policy Committee will publish its updated assessment on 7 July, so I won't front-run that. But I will note that in April, we concluded that while to date AI infrastructure investment had thus far mostly been financed by large technology companies' cash flows and equity, debt financing was rising rapidly, and in some new and complex ways. We therefore judged that the financial stability consequences of any fall in AI-related asset prices could well increase.^[8]

Financial stability implications of an agentic system, and how central banks should respond

Cyber capabilities

My most proximate financial stability concern however is what the UK Government's AI Security Institute has described as a step change in agentic AI's cyber capabilities.^[9]

In the hands of defenders, these tools strengthen cyber resilience. But, in malicious hands, they materially increase the chance of attacks that could harm financial stability. And so, the immediate challenge is to maintain the advantage of cyber defenders as these capabilities continue to evolve, particularly in an increasingly severe cyber threat environment including from geopolitics and ransomware.

Cyber resilience has been a priority for the Bank of England and our G7 colleagues for over a decade^[10], and the Bank has worked closely with other UK authorities, Government and the National Cyber Security Centre. In May, the Bank, FCA and HM Treasury set out practical

steps for financial institutions to prepare for these frontier AI-related cyber risks.^[11] Getting the fundamentals of cyber hygiene right is more important than ever.

But this gives only limited comfort. AI is uncovering many vulnerabilities to patch, including as global software providers and key UK financial institutions use the latest models. For financial stability, patching must happen quickly – across the financial sector, key third-party technology providers, and wider national infrastructure such as energy and telecoms. That is no small task.

It's also an urgent one. As the heads of the Five Eyes cyber security agencies said last week, “the timeline is not years, it is months”.^[12] Open source models may lag the most advanced closed models by only four to eight months, and their safeguards can be breached.^[13] If patches are released but not swiftly implemented, malicious actors can reverse engineer the vulnerabilities they address.

For central banks, this increases the onus on a system-wide approach to operational resilience.^[14] Stronger firm defences are essential, but not enough. Firms must also respond to and recover from disruption quickly – whether from successful cyber attacks or from patching itself causing operational disruption, as the faulty CrowdStrike update did in 2024.^[15]

Authorities should place higher likelihood on simultaneous multi-firm disruption, enhance tools to stress test that impact ex ante^[16], and strengthen mechanisms with industry to coordinate ex post.^[17]

We should also consider whether regulated firms need enhanced recovery options for core systems. In the event of financial failure, we have deposit insurance so customers can still access their money, and bail in-able debt so the largest firms can continue to provide services. In a cyber context, do we need systems that allow one institution to pick up another's basic functions during disruption, as in Ukraine's “Power Banking” programme launched in 2022?^[18] Or requirements for key firms to have completely separate fail-over capabilities or to be able to rebuild compromised core systems quickly from “bare metal”? These are big questions, but they are the right ones.

Agentic trading in financial markets

Turning to agentic trading in financial markets, evidence suggests that for now, trading firms mostly use autonomous AI for lower-risk operational tasks, such as research.^[19] But that could change quickly.

If AI agents respond similarly to the same prompts or triggers, they could amplify volatility in stress – especially if their objectives drift from original goals or public policy objectives, in a manifestation of the misalignment problem that can arise with some AI models.^[20]

The financial stability question therefore broadens from whether firms can use models well to whether the system can also observe and contain their resulting behaviours.

We have experience of structured, multi-round stress simulations focused on how financial market participants behave as they learn of each other's actions^[21]. AI increases the importance of such work - and it should transform how we do it. We are experimenting with the BIS Innovation Hub and the Bundesbank on simulation methods to understand which aspects of agent design could drive herding behaviour.^[22]

That work can also explore mitigants: whether markets using AI agents are resilient enough; whether agents' objective functions could incorporate public policy objectives; and whether guardrails are needed, analogous to circuit breakers or kill switches that would limit or stop trading market-wide if faulty AI models cause market meltdown.

More broadly, AI should transform central banks' monitoring and risk assessment. Agentic technology could combine a wide range of qualitative and quantitative data to generate insights and scenarios for how the current conjuncture could go wrong, potentially using a 'digital twin' of the financial system to simulate interactions between participants.^[23]

Agentic payments and commerce

Turning lastly to e-commerce and payments, consumers and merchants have so far mainly used AI agents in recommendation mode, with humans still executing transactions. But as agentic technology advances, tech firms, payment systems and merchants are working to automate that final step (such that I could ask an AI agent to 'book my holiday', 'refill my fridge' or 'refresh my wardrobe').

That matters for central banks. In the UK, the Bank is leading a public-private partnership to design the next generation of retail payments infrastructure, ensuring it can support agentic payment journeys and expand choice and functionality for users.^[24]

But the biggest issues are likely to be for regulation^[25] and industry standards: how users securely give consent and authorisation to agents, especially for multiple transactions; how disputes are settled and liability assigned for erroneous or fraudulent transactions; and how authorities avoid fragmentation and walled gardens as AI firms and payment systems all develop protocols for agents to interact with themselves and merchants.

What these two examples – agentic commerce and agentic trading – both highlight is that, as AI capabilities increase, we must keep asking whether existing, technology-agnostic regulatory frameworks remain sufficient. Our frameworks were not built to contemplate autonomous agents, and relying on a human in the loop for all agent actions is unlikely to be realistic. More sophisticated governance and accountability frameworks may be needed.^[26]

Conclusion

To conclude, AI capability is accelerating and becoming increasingly agentic. The financial system is likely to evolve into one that operates more autonomously, at scale and speed. The transition is uncertain and will bring risks of its own to monitor. And a more agentic system means central banks must think hard not only about how finance harnesses the opportunities and mitigates the risks, but also about how new technological possibilities should change how we do our own job.

We need policy frameworks, internal skills and institutional structures ready for more frequent technology surprises. Cyber risks, agentic trading and agentic payments and commerce are the applications on my radar today. But that list could look different in a year, if not sooner.

Scenario analysis, and how we prepare now in light of it, will be key. What if the next surprise puts the latest capabilities in bad actors' hands, or models develop in ways that are harder to evaluate and control, given evidence that some behave differently in testing compared to real-life scenarios?^[27]

Cooperation with other parts of the official sector is vital – for example, drawing on AI security institutes' model testing and strengthening resilience through domestic AI capabilities and capacity.

The same applies internationally, including in the Financial Stability Board^[28] and the G7.^[29] The global financial crisis underlined how interconnected finance is – the FSB came out of that.

The challenge of AI is similar. The impact of new AI capabilities can spread across borders through common technology dependencies, globally systemic financial institutions and market infrastructure.

We shouldn't wait for a crisis to build the cooperation we need. With other central banks, regulators and finance ministries, the Bank of England will do our part to ensure responsible AI adoption and, through it, sustainable economic growth.

I would like to thank Kimberley Moran and Michael Yoganayagam for their assistance in drafting these remarks, and Owen Lock, Irina Mnoghitnei and Andy Walters for their helpful input.

-
1. The Bank's considerable programme of work on this front is set out in this 1 April 2026 [letter](#) to Parliament's Treasury Committee. The Bank and FCA first [surveyed](#) use of AI and machine learning in the financial system in 2019.
 2. As discussed by Andrew Bailey in his speech in May: [Can AI make Cutlery? - speech by Andrew Bailey | Bank of England](#)

3. [Measuring AI Ability to Complete Long Tasks - METR](#)
4. [Time Horizon 1.1 - METR](#)
5. [How fast is autonomous AI cyber capability advancing? | AISI Work](#)
6. Figure 1, Ibid.
7. [Introducing OpenAI o1 | OpenAI](#)
8. [Financial Policy Committee Record – April 2026 | Bank of England](#)
9. [Our evaluation of Claude Mythos Preview’s cyber capabilities | AISI Work](#)
10. [G7 Fundamental Elements of Cybersecurity for the Financial Sector - GOV.UK](#)
11. [The Bank, FCA and HM Treasury joint statement on Frontier AI models and cyber resilience | Bank of England](#)
12. [Five Eyes cyber security agencies statement, The AI shift in cyber risk: why leaders must act now](#)
13. [Frontier AI Trends Report by The AI Security Institute \(AISI\)](#) [Managing risks from increasingly capable open-weight AI systems | AISI Work](#)
14. [Financial Stability in Focus: The FPC’s macroprudential approach to operational resilience | Bank of England, March 2024.](#)
15. [IT Disruptions from CrowdStrike’s Update: Frequently Asked Questions | Congress.gov | Library of Congress](#)
16. [Cyber and Operational Resilience Stress Test \(CORST\) | Bank of England](#)
17. [About CMORG | Cross Market Operational Resilience Group](#)
18. [National Bank of Ukraine announcement, December 2022](#) Andrew Bailey also praised this work during his visit to Kyiv in 2025: [Central Banking in extreme adversity - speech by Andrew Bailey | Bank of England](#)
19. [FMSB Spotlight Review offers a practitioners’ view of how AI is being used in markets | Financial Markets Standards Board](#)
20. [Investigating models for misalignment | AISI Work](#)
21. [System-wide exploratory scenario | Bank of England](#)
22. [Project Logos: observing the behaviour of LLM-based agents in a simulated financial market environment](#)
23. [Cybernetic Financial Stability Monitoring - Mike Hsu](#)
24. [The Retail Payments Infrastructure Board launches consultation on the next generation UK payments infrastructure | Bank of England, 25 June 2026](#)
25. HMT intends in the UK to explore through a forthcoming consultation how the regulation of payments services should adapt to payments conducted by AI agents: [UK fintech backed to embrace future payments technology - GOV.UK](#)
26. We are exploring these issues in the Bank-FCA AI Consortium, comprised of experts from the financial industry, tech sector and academia: [Artificial Intelligence Consortium | Bank of England.](#)
27. [Will it become harder to oversee AI systems? | AISI Work](#)

28. **FSB consults on sound practices for the responsible adoption of artificial intelligence (AI) - Financial Stability Board** 10 June 2026

29. G7 Leaders agreed on 17 June 2026: *“In light of the rapid advancement of capabilities of frontier artificial intelligence models, we ask our Ministers of Finance and Central Bank governors, in coordination with financial supervisors and representatives of global financial institutions and tech companies, to further discuss emerging opportunities and potential risks arising from artificial intelligence, including in the financial sector, while considering implications for productivity and labour markets. We also ask the G7 cyber expert group to, as appropriate, enhance information sharing and identify best practices, in light of the recent developments regarding frontier artificial intelligence models. We also encourage further dialogue between cybersecurity agencies and relevant institutions in existing G7 groups.”* **Leaders' statement for a more balanced durable resilient growth. | G7 Évian 2026**



Sarah Breeden

Deputy Governor, Financial Stability