FUTUREMATTERS™
CENTRE OF EXCELLENCE

GFTN
GLOBAL FINANCE & TECHNOLOGY NETWORK

Implications of AI for Financial Stability: Global Perspectives

Aug 2025

Contents

About

Deep Future Analytics (DFA) is the result of 30 years of research and experience in credit risk analytics, in all its many aspects. With a team of 25 expert data scientists, technologists, and economists, DFA is led by Dr. Joseph Breeden, a pioneer in the field of risk modeling. Together, we equip financial institutions with actionable intelligence to drive smarter, more strategic decisions.

For more information, visit https://www.deepfutureanalytics.com/en

The Global Finance & Technology Network (GFTN) is a Singapore-headquartered organisation that leverages technology and innovation to create more efficient, resilient, and inclusive financial systems through global collaboration. GFTN hosts a worldwide network of forums (including its flagship event, the Singapore FinTech Festival); advises governments and companies on policies and the development of digital ecosystems and innovation within the financial sector; offers digital infrastructure solutions; and plans to invest in financial technology startups through its upcoming venture fund, with a focus on inclusion and sustainability.

For more information, visit www.gftn.co

The GFTN Forum, Japan - formerly known as Japan FinTech Festival - is part of the Japan Financial Services Agency's Japan FinTech Week (JFW). This Forum highlights the dynamic international ecosystem within JFW, embracing the bold theme of "Building Financial Corridors Worldwide" while supporting the domestic ecosystem.

For more information, visit https://gftn.co/global-forums/gftn-forum-japan

GFTN Insights is a year-round series that connects senior officials and industry leaders across continents to address key financial and tech challenges. Held under Chatham House rules, these dialogues foster partnership, culminating in the two-day Insights Forum in Singapore.

For more information, visit https://gftn.co/programmes/gftn-insights

1 Introduction

During Japan FinTech Week in Tokyo, a closed-door roundtable was convened as part of the GFTN Forum, Japan to discuss the regulatory, technological, and systemic implications of artificial intelligence (AI) in financial services. Participants included representatives from global financial regulators, central banks, academia, venture investment, and the private sector. Since Chatham House Rules were being observed, all comments are anonymous, although many of the points made are consistent with public statements of the speakers.

2 Framing the AI Regulatory Challenge: The Trilemma

The session opened with a theoretical framework describing the regulatory trilemma for AI: the challenge of achieving innovation, financial integrity, and regulatory clarity simultaneously. Historically, jurisdictions have attained at most two of these goals at a time. This trilemma becomes even more pronounced for AI, which adapts rapidly and unpredictably.

Participants emphasised the difficulty in balancing a permissive environment conducive to innovation with the need for robust oversight to maintain market integrity. Regulatory clarity, while desirable, often comes at the cost of innovation or integrity, especially in the early phases of technological adoption.

3 Jurisdictional Approaches to AI Oversight

3.1 European Union: Codifying risk-based regulation

The EU AI Act[1] was the first comprehensive legal framework for AI. Its key feature is a risk-tiered classification: unacceptable, high, limited, and minimal risk. High-risk systems, such as those influencing credit decisions, require a Fundamental Rights Impact Assessment and registration in an EU database. Providers, deployers, importers, and distributors all have defined obligations under the Act.

The legislation prohibits social scoring systems and manipulative AI practices. Specific examples were given, such as an AI system used in loan pre-screening being permitted if a human makes the final decision. The speaker highlighted the severity of penalties: up to 7% of global annual turnover or €35 million for the most serious violations, with reduced caps for SMEs and startups.
Note that the EU AI Act has been amended as to the interpretation of "high risk" for financial institutions. Some initial interpretations suggested that data-driven machine learning models as used in credit scoring could fall under these restrictions, but this has been scaled back to allow a continuation of applications of machine learning.

3.2 United States: Principles-based and sector-specific

The U.S. approach remains principles-based, relying on pre-existing frameworks such as model risk management guidance (SR 11-7), consumer protection laws, and internal controls.[2] Regulators are engaging with supervised entities through examinations and interagency requests for information.

A key theme was the adaptability of existing risk categories (e.g., cyber, operational, model, and consumer risk) to AI, particularly traditional machine learning. While current practices accommodate incremental AI evolution, there remains concern about emergent risks that may require new regulatory tools.

Although the principles may remain unchanged, many organisations are considering whether the best practice application of those principles needs to be adapted for the unique features of generative AI. For example, although validation must still be performed, model monitoring has heightened importance. From the first day of deployment, a generative AI model may exhibit unexpected behaviours. Unlike with traditional data-driven models, validating generative AI models does not create a reliability period before retesting is required.

3.3 Singapore: Iterative maturity through collaborative development

Singapore has developed the FEAT principles (Fairness, Ethics, Accountability, Transparency)[3] and subsequently Project MindForge[4], which engages industry actors across

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139, and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797, and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance).

[2] Brainard, L. (2018, November 13). What are we learning about artificial intelligence in financial services? Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/newsevents/speech/brainard20181113a.htm

[3] Monetary Authority of Singapore. (2018, November 12). MAS unveils principles to promote fairness, ethics, accountability and transparency in use of AI and data analytics by financial institutions. Retrieved from https://www.mas.gov.sg/news/media-releases/2018/mas-unveils-principles-to-promote-fairness-ethics-accountability-and-transparency-in-use-of-ai-and-data-analytics-by-financial-institutions

[4] Monetary Authority of Singapore. (2023, November 15). MAS partners industry to develop generative AI risk framework for the financial sector. Retrieved from https://www.mas.gov.sg/news/media-releases/2023/mas-partners-industry-to-develop-generative-ai-risk-framework-for-the-financial-sector

financial services. This iterative engagement strategy has enabled the co-creation of detailed use case guidance. Singapore's regulators stressed the importance of starting small, observing outcomes, and iterating. Their regulators also collaborate closely with national agencies overseeing AI and data governance, ensuring cohesion across regulatory domains.

3.4 Japan: Encouraging safe adoption via dialogue

Japan published a discussion paper highlighting AI use cases, associated challenges, and existing practices, with the goal of stimulating industry-regulator dialogue.[5] Rather than immediately establishing hard rules, Japan emphasised clarifying regulatory expectations as a first step. Formal action would be considered only if clear regulatory gaps emerge.

4 Key Use Cases of AI in Finance

4.1 Internal operations and productivity tools

Across jurisdictions, participants noted that most firms currently use AI for internal productivity: document summarisation, email drafting, call routing, and research. These use cases were described as generally low risk and serve as entry points for adoption. However, more recent discussions are highlighting risks that may go beyond the initial perception.

At one institution that was deploying an internal, enterprise chatbot, HR employees immediately asked if they could screen resumes for similarity to job postings. Legal gave a hard "No". What if a loan officer asks for background information on a company seeking a loan (potentially allowed) and then asks if the loan should be given (definitely disallowed). A prohibition against "decision support" may be adopted as a corporate policy, but are guidelines and training sufficient to assure low-risk use of internal AI systems? Increasingly, institutions are recognising that such interactions may require monitoring, just as the IT department monitors inappropriate website access.

Generative AI models, such as large language models (LLMs), are being tested for regulatory compliance automation, early-stage drafting of communications, and knowledge base querying. However, hallucinations and a lack of explainability remain barriers to full automation.

4.2 Credit risk and decisioning

AI in credit scoring remains a focal point for regulatory attention. It was widely acknowledged that AI is increasingly embedded in credit decisions, either directly or as a pre-screening filter, but such data-driven machine learning models should be distinguished from unbounded generative AI models. While full automation of credit approvals by AI systems is prohibited in some jurisdictions, hybrid systems where humans retain final authority are more common.

Bias in training data and feature selection, especially involving protected characteristics like zip code or education history, was cited as a significant concern. Without proper controls, historical discrimination can be perpetuated by AI. Discussions around bias are finally beginning to recognise that limiting inputs is insufficient to avoid discriminatory outcomes. Highly nonlinear machine learning algorithms may discover patterns in default that correlate to protected class status, because societal inequities make this an embedded aspect of the training data. Solutions are still being explored, but hiding the applicant's status from model developers only ensures that bias testing cannot be performed, not that bias will be prevented.

4.3 Fraud detection and AML compliance

AI is also being deployed in anti-fraud and anti-money laundering (AML) contexts, particularly in transaction monitoring and anomaly detection. Participants noted that while these tools are valuable, adversarial actors are also using AI to create deepfakes and circumvent controls. Because of the speed and sophistication of adversarial AI techniques, institutions may have no choice but to deploy equally adaptive AI systems.

[5] Financial Services Agency of Japan. (2025, March 4). Preliminary Discussion Points for Promoting the Sound Utilization of AI in the Financial Sector (Version 1.0). Retrieved from https://www.fsa.go.jp/en/news/2025/20250304/aidp.html

4.4 Sentiment analysis and market forecasting

Sentiment analysis, often paired with trading algorithms, represents a more speculative use case. These stack multiple innovations (e.g., LLMs with market infrastructure) and can create layers of opacity. Their use raises questions about transparency, accountability, and potential herding behavior.

5 Risk Considerations and Mitigating Controls

5.1 Traditional risk categories reapplied to AI

Most participants agreed that AI amplifies existing risks rather than introducing wholly new ones. These include:

- **Cyber risk:** Both in terms of attack vectors and defense mechanisms.
- **Operational risk:** Due to misconfiguration, failure to update models, or inadequate governance.
- **Model risk:** Especially in blackbox systems and models with complex training pipelines.
- **Consumer protection:** Including bias, explainability, and rights of appeal.

Deploying generative AI is so different from previous statistical and machine learning approaches, and the applications so much more varied, that the greatest risk is not knowing which risks might apply. Applications such as the internal chatbot described above may seem low risk until users invent new, unapproved uses.

Rather than focusing on how best to add AI models and tools to the existing model inventory, model risk managers need to incorporate their own lack of knowledge. These "unknown unknowns" may invalidate the normal risk level assessments. As such, even seemingly low-risk applications require vigilance upon initial deployment to make sure that they are being used as anticipated. An annual review for "low risk" AI deployments may come far too late.

5.2 Third-party dependencies and concentration risk

A recurrent concern was the increasing dependency on third-party AI vendors, particularly major cloud and model providers. This raises systemic risk via concentration and challenges for regulatory oversight when services are outsourced.

While regulatory responsibility remains with the supervised institution, the opacity and market power of upstream vendors complicate assurance. Divergent regulatory frameworks across countries further hinder harmonised governance.

5.3 Lack of data and model transparency in pretrained models

Many AI tools used in financial services are pretrained on undisclosed data. This undermines firms' ability to verify data quality, provenance, and suitability. Without control over model development, firms struggle to conduct proper validation.

An inability to establish data provenance and rights for use in model development also potentially carries legal risk for institutions using these tools. Beyond the need for model risk management to verify these rights, legal contracts should be amended to transfer any associated risks to the vendor.

5.4 Shifting emphasis from validation to monitoring

Much is being written about validating generative AI models, but these models violate the core assumptions of model validation. With traditional models, the input and output spaces can be fully explored to verify model robustness, or tail events in input or output data can be explicitly limited.

Generative AI models can be explored, but no amount of validation testing can assure that the model will perform as expected once deployed. On day 1, a user may take unexpected actions, such as the discussion around "jail breaks", that take the model outside the validation bounds. Validation is still necessary, but cannot be sufficient. Instead, model risk management will need to shift to an emphasis on model monitoring.[6]

[6] Breeden, J.L. 2025. Effective Generative AI Model Risk Management. ResearchGate. https://www.researchgate.net/publication/390760580_Effective_Generative_AI_Model_Risk_Management

6 Financial Stability Considerations

6.1 Systemic concentration and herding behavior

Concentration risk among AI service providers could reduce systemic resilience. If most firms rely on the same models or data providers, common failures or flawed updates could cascade through the system.

Concerns about herding behavior are rising: if AI systems interpret signals similarly or rely on shared data, market moves may become synchronised. However, some argue that demand for differentiated alpha will push firms toward diverse models.

6.2 Adversarial use of AI

AI is also augmenting financial crime. Fraudsters, cybercriminals, and hostile actors are using AI to scale attacks, generate synthetic identities, or subvert biometric verification. The asymmetry between defensive and offensive capabilities may widen in the short term.

6.3 Data gaps and regulatory blind spots

Regulators highlighted a lack of structured information on AI usage within supervised firms. Surveys and voluntary disclosures provide partial insights, but systematic data is lacking. This impedes macroprudential analysis of AI-driven risks. This could be addressed in part via compliance monitoring for AI systems in production.

6.4 Long-term impacts on market structure

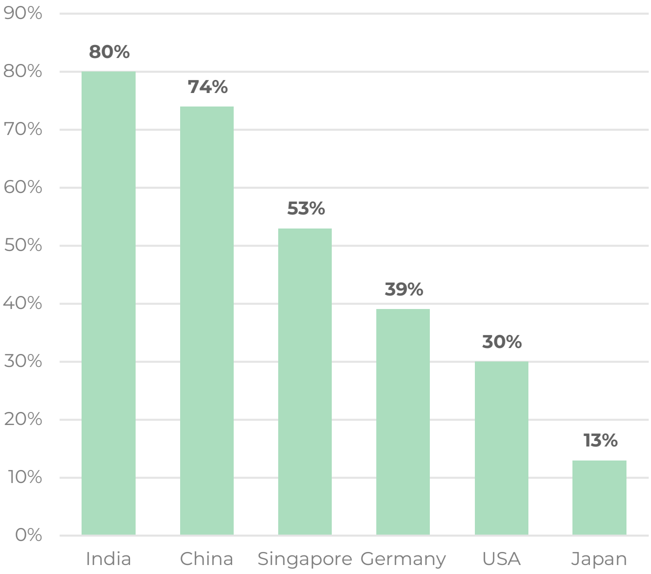
Participants discussed potential long-term shifts, including job displacement, changes in firm competitiveness, and effects on capital allocation. If AI enables massive productivity gains, it could widen gaps between early adopters and laggards.

One speaker warned that AI may be beneficial over time, but abrupt changes in adoption could cause destabilisation. For example, when interest rates change rapidly, this shocks markets and consumers, who are reacting to the rate of change rather than the absolute level of interest rates. This led directly to the 2009 Global Financial Crisis and the 2023 collapse of Silicon Valley Bank. When considering the impact of AI, it was suggested that what matters most is not the level of AI adoption and economic displacement but the rate of change during the transition.

7 Building Trust and Capabilities

Trust was a central theme of the discussions. One participant shared excerpts from a recent government study about societal reluctance toward AI adoption, measuring public trust in AI under current regulation. India had 80% trust in AI. China was 74%, Singapore 53%, Germany 39%, the US 30%, and Japanese trust was measured at just 13%. Without trust, AI adoption will be limited and business opportunities missed.

Public Trust in AI



Source: GFTN Analysis

Firms are hesitant to deploy AI in client-facing functions due to fear of regulatory censure or reputational risk. Regulators, for their part, require assurance frameworks that are still in development.

7.1 AI-augmented monitoring

Human-in-the-loop (HITL) oversight is a commonly proposed solution, but psychologists have long known that humans perform poorly when tasked with finding rare failures in high-volume systems. Instead, vendors such as SAS and Deep Future Analytics have introduced LLM tools as a second line of defense monitoring of frontline AI systems.

Questionable communications or actions are flagged for review by Model Risk Management (MRM) personnel, making HITL oversight feasible. Failures may be relative to regulatory, ethical, or business requirements.
The greatest advantage of such monitoring systems may be in the creation of compliance performance metrics. Where speakers noted a lack of performance metrics, continuously tracking compliance metrics could significantly improve trust in the systems.

7.2 Applying Model Risk Management to Humans

AI models are being deployed in areas previously staffed by humans. As metrics are developed for monitoring AI accuracy and compliance, the greatest challenge may be in establishing a benchmark. AI systems will not be perfect, so how good is good enough?

This problem was encountered in the 70s and 80s with the introduction of credit scores to lending. Credit score performance was acceptable so long as it outperformed what human loan officers had been doing judgmentally. We need the same comparison for AI deployments, meaning that we need to apply model risk management principles to the human service providers who are being replaced or supplemented by AI. In fact, applying monitoring to a mixed environment of human and AI agents is exactly what is needed to establish trust in AI.

8 Conclusion: The Need for Collaborative Adaptation

Despite varying national strategies, participants broadly agreed on several principles:

- Begin with low-risk use cases, but do not avoid engaging high-risk domains.
- Encourage dialogue between regulators, industry, and technologists to refine risk controls.
- Develop supervisory capacity through AI-enabled tools.
- Monitor systemic risks associated with concentration, herding, and opaque outsourcing.
- Strengthen cross-border cooperation to align on data, oversight, and best practices.

As AI becomes more embedded in financial systems, the balance between innovation, integrity, and clarity will require continuous recalibration. The roundtable underscored that while there are no universal solutions, sharing experiences and co-developing guidance is an essential foundation for global stability.

Authors

GFTN Research & Advisory

Dr. Joseph L. Breeden
Chief Executive Officer, Deep Future Analytics

Contributors

Akanksha Rath
Head of Capacity Building and Learning Initiatives

Production

Sachin Kharchane
Graphic Designer

Global Finance &
Technology Network (GFTN)

6 Battery Road, #28-01, Singapore 049909
gftn.co | hello@gftn.com

This document is published by Global Finance & Technology Network Limited (GFTN) as part of its FutureMatters insights platform. The findings, interpretations, and conclusions presented in GFTN Reports reflect the views of the author(s) and do not necessarily represent those of GFTN, its Board, management, stakeholders, or any individual participant and their respective organisations.