

Risks: models, bias, herding and fraud

*Dr Alexander Ressos, TEP
Attorney-at-Law (Switzerland), Solicitor (Hong Kong),
Solicitor (England & Wales)*

The risks of artificial intelligence in financial services fall into several categories that supervisors and the literature have identified. Model risk is the oldest. Every model simplifies reality and may fail when conditions change. Machine learning models are particularly exposed because they learn patterns from historical data that may not hold in the future. The complexity of large models makes it difficult to validate them in the way that banks have validated statistical models since the 2000s.¹

Bias and discrimination form a second category. A credit model trained on historical decisions may reproduce the patterns of those decisions, including any disadvantage for certain groups, even when it does not use protected characteristics directly. Swiss law contains no general prohibition of discrimination between private parties, although the protection of personality under Article 28 of the Civil Code and specific statutes offer some protection. FINMA's guidance lists equal treatment among its areas of attention, while the EU AI Act addresses the issue through its requirements on data governance for high-risk systems.² Explainability is closely related. When a model cannot explain why it rejected a loan application or flagged a transaction, the client cannot challenge the decision and the institution cannot demonstrate that it acted lawfully.

Financial stability is the concern of the institutions in Basel. If many market participants use similar models trained on similar data, their decisions may become correlated, which could amplify market movements and reduce the diversity of views that keeps markets liquid. Gary Gensler, before he became chair of the United States Securities and Exchange Commission, and Lily Bailey described this herding risk in a paper of 2020, arguing that deep learning could increase systemic risk through uniformity, network interconnectedness and regulatory gaps.³ The Financial Stability Board

added concentration among a few providers of models, cloud services and chips, which creates single points of failure for the financial system.⁴

Cyber risk and fraud have grown with the technology. Attackers use language models to write convincing phishing messages in any language, including Swiss German, and to identify weaknesses in software. Voice cloning and video deepfakes allow criminals to impersonate executives, clients or relatives. In a widely reported case in 2024, an employee of the engineering firm Arup in Hong Kong transferred about USD 25 million after a video conference in which the other participants, including the apparent chief financial officer, were generated by artificial intelligence.⁵ The National Cyber Security Centre, which became the Federal Office for Cyber Security in 2024, has reported a growing number of such attempts in Switzerland.⁶ For banks and their clients, the consequence is that a voice or a face on a screen no longer proves identity, so that payment instructions require independent verification through a separate channel.

1 Notes

1. On model risk management, see the supervisory guidance of the Federal Reserve and the Office of the Comptroller of the Currency, SR 11-7, Guidance on Model Risk Management, April 2011, which has influenced practice internationally. To be verified. ↩
2. Art. 28 of the Swiss Civil Code (ZGB, SR 210); FINMA Guidance 08/2024; Art. 10 of Regulation (EU) 2024/1689 on data and data governance. To be verified. ↩
3. Gary Gensler and Lily Bailey, "Deep Learning and Financial Stability", working paper, MIT, November 2020. To be verified. ↩
4. Financial Stability Board, The Financial Stability Implications of Artificial Intelligence, 14 November 2024, sections on third-party dependencies and market correlations. To be verified. ↩
5. Reports of the Hong Kong Police and international press reports of February and May 2024 on the deepfake fraud against Arup's Hong Kong office. Figures to be verified. ↩

6. Federal Office for Cyber Security (BACS), semi-annual reports and weekly reviews on fraud and deepfakes since 2024. Status to be verified.
↩

For general information only. This article does not constitute legal or tax advice, does not replace advice on an individual case and creates no lawyer-client relationship. It reflects the legal position as of September 2026.

Online version: helveticaregister.com/basel-risks-models-bias-herding-and-fraud