

新加坡：MAS 與業界共同推動 AI 代理安全防護機制（7/3）

- 新加坡金融管理局（MAS）與主要金融機構及金融科技業者，共同發布《執行階段代理式金融安全防護機制（Safeguards for Agentic Finance at Runtime, SAFR）》^{註1} 白皮書，提出一套由產業共同研擬之架構，以協助金融服務中的人工智慧代理（AI Agents）安全、可靠且符合資訊安全要求地執行各項金融任務。SAFR 係於 MAS 推動之 BuildFin.ai ^{註2} 倡議下之進一步發展，旨在協助金融業負責任地開發及推展人工智慧解決方案。相關重點如下：

- 建立 AI Agents 安全防護機制

隨著 AI Agents 逐漸能自主且快速執行金融服務相關工作，其執行速度已超出人員可即時介入之範圍，因此金融機構須建立即時安全防護機制，確保 AI Agents 之行為符合預先設定之指令及風險界限。SAFR 架構透過治理檢核機制，在人工智慧代理執行任務前，先針對其擬採取之行動進行驗證及紀錄。

- 強化 AI 治理機制

SAFR 以 MAS 「Project Mindforge」^{註3} 人工智慧風險管理工具包為基礎，著重於如何將安全防護機制落實於 AI Agents 執行行動之階段。白皮書並提出未來發展方向，說明如何依政策限制執行（policy-bound execution）、即時驗證、可稽核性及互通性（interoperability）等機制整合至系統運作流程，使金融機構能一致且可信賴地推展 AI Agents。

- 擴大金融業應用情境

目前業者已將 SAFR 架構應用於支付與資金管理、財富管理及投資顧問，以及客戶服務等情境，藉由 AI Agents 執行例行作業、文件審閱及客戶分析，以提升營運效率、法令遵循及客戶服務品質。

➤ 持續推動產業合作

MAS 邀請產業夥伴加入 BuildFin.ai 工作小組，共同參與 SAFR 後續版本之研擬與精進。另近期宣布成立之未來金融研究所（Future of Finance Institute, FFI）^{註4}，將透過推動產業試辦計畫及監理沙盒實驗，協助金融機構測試及推展符合 SAFR 架構之解決方案。

註1：《執行階段代理式金融安全防護機制》白皮書相關內容詳見：
<https://www.mas.gov.sg/publications/monographs-or-information-paper/2026/safeguards-for-agentic-finance-at-runtime>

註2：「BuildFin.ai」為 MAS 於 2025 年推出之產業合作倡議，透過串聯金融機構、金融科技業者及科技公司，共同開發、測試及驗證 AI 解決方案，並建立相關技術標準、最佳實務及治理機制，以促進 AI 於金融服務之安全、可信賴且負責任的應用。

註3：「Project Mindforge」為 MAS 推動之 AI 治理倡議，提供 AI 風險管理工具包（AI Risk Management Toolkit），協助金融機構建立 AI 治理及風險控管機制，並作為 SAFR 架構發展之基礎。

註4：MAS 於 2026 年 6 月宣布成立之未來金融研究所，詳細內容請參閱 2026 年 6 月國際證券期貨市場動態（編號：260616）。

資料來源：[新加坡金融管理局 MAS](#)