金融穩定委員會（FSB）「人工智慧對金融穩定之影響」摘要報告

本公司國關室摘譯

壹、 摘要

本報告概述近期人工智慧（AI）技術發展趨勢，並評估其對金融穩定之潛在影響。本報告回顧金融穩定委員會（FSB）於 2017 年發布之 AI 應用報告，檢視最新技術進展，探討 AI 在金融領域的應用場景、潛在效益及相關脆弱性等。

儘管金融業早已開始使用 AI 工具，近年來其應用範圍已顯著擴展。2017 年 FSB 報告曾列舉 AI 於金融體系中的重要應用，包括客戶端應用（如自動化服務、評估授信品質）、營運端應用（如資本優化、風險管理）、法遵科技（RegTech）及監理科技（SupTech）等。自報告發布以來，AI 應用

本文係中央存保公司摘譯 2024 年 11 月 14 日發表於金融穩定委員會（FSB）之「人工智慧對金融穩定之影響」（The Financial Stability Implications of Artificial Intelligence），非 FSB 官方翻譯。本篇文章無償取材自金融穩定委員會網站（www.fsb.org/），本文中譯內容如與原文有歧義之處，概以原文為準，原文網址連結如下：https://www.fsb.org/uploads/P14112024.pdf

持續拓展，並隨生成式 AI（Generative AI, GenAI）[註1]技術進步而迅速發展。金融機構及其供應商正探索生成式 AI 在客戶服務、詐騙檢測、市場分析、文件處理、資訊檢索及軟體開發等領域的應用潛力。AI 雖能為金融市場、金融機構及其客戶帶來效益，但同時可能加劇金融業的脆弱性，如第三方依賴、市場相關性、網路安全、詐騙及不實資訊傳播等風險。

貳、 AI 發展

自 2017 年 FSB 報告發布以來，AI 快速發展，其中以深度學習（deep learning）[註2]、大數據、電腦運算能力及生成式 AI 尤為顯著。2022 年 11 月 ChatGPT 推出，進一步引發市場對 AI 的關注，AI 相關的就業人數及專利數量亦持續上升。2017 年 FSB 報告分析 AI 在金融服務領域應用的供需雙方誘因。隨著技術進步，供給面影響愈發重要；雖需求面仍是推動 AI 應用的重要誘因，但自 2017 年以來未有顯著變化。

一、 供給面誘因

(一)技術發展

生成式 AI 結合軟硬體進步，提升 AI 的吸引力。深度學習模型強化 AI 處理非結構化數據（unstructured data）[註3]的能力，透過嵌入技術（embeddings）[註4]實現高效數據處理；同時，變換器架構（transformer architecture）[註5]的發展徹底革新自然語言處理（NLP）[註6]，為生成式 AI 與大型語言模型（LLMs）[註7]的發展奠定基礎。LLMs 將人類與電腦的互動方式從依賴程式碼轉變為使用人類習慣之文字或語音，使技術更易於普及；硬體方面，圖形處理器（GPU）[註8]大規模的應用整合及其計算能力之提升，使處理大型數據及應用複雜模型的成本得以降低。

根據國際貨幣基金（IMF）估計，金融服務業在 AI 系統軟硬體

及服務上的投資預期將從 2023 年的 1,660 億美元增至 2027 年的 4,000 億美元。

儘管技術驅動 AI 普及，人力資源的發展卻未能同步跟上。金融機構對於 AI 人才的需求日益增加；然而，具開發 AI 模型（尤其是生成式 AI）能力的專業人才卻十分稀缺且昂貴。

(二)數據相關發展

隨著客戶與金融機構之間的數位互動日益增加，數據的可用性及規模大幅增長，進一步加速 AI 應用。由於技術發展，金融機構除能處理傳統量化數據外，亦能分析圖片及影片等非結構化數據。同時，許多機器學習（Machine Learning, ML）（註 9）模型（特別是生成式 AI 模型）需要大量數據進行訓練。模型對數據的需求，可能導致高品質數據枯竭。為因應數據不足、隱私權及智慧財產權等挑戰，使用合成數據（synthetic data）（註 10）成為解決方案之一。然而，合成數據本身也伴隨著風險與限制，包括生成數據的品質管理等。

(三)商業模式發展

近年來，雲端運算與預訓練 AI 模型成為兩大重要商業模式發展方向：

1.雲端運算

雲端技術提供靈活的計算及資料儲存資源，近年來，部分金融機構已將其應用於核心業務，並提升對第三方 AI 服務供應商的接受度。由於運算基礎設施及模型訓練成本上升，許多機構選擇依賴雲端供應商（特別是生成式 AI 模型）。

2.預訓練 AI 模型（pre-trained AI models）（註 11）

預訓練 AI 模型降低金融機構對專業 AI 技術開發的需求，促進 AI 的普及應用。然而，模型的解釋性（explainability）（註 12）與

問責性（accountability）不足仍構成重要挑戰。

此外，開源（Open Source）[註 13] 模型日益興盛，預訓練 AI 模型主要分為開源與封閉程式碼兩種。儘管封閉源程式碼模型過去在多數測試中表現優於開源模型，但近年來開源模型性能顯著提升，部分開源 AI 模型已在多數任務中與封閉模型表現相當。開源模型通常免費，且可根據機構需求高度客製化，有助於降低 AI 應用的成本。2023 年，新發布的基礎模型中 66% 為開源，相較 2021 年的 33% 大幅增加。

二、 需求面誘因

(一)增強盈利能力

透過增加營運收入、降低成本、強化風險管理及提升生產力來增強盈利能力，至今仍是推動金融機構部署 AI 的重要因素。

1. 增加營運收入：幫助金融機構提供客製化的產品與服務，簡化客戶開戶流程，降低客戶流失率。
2. 降低成本：將部分人工作業自動化，節省人力成本。
3. 強化風險管理：預測現金流、貸款拖欠率及超額損失等。
4. 提升生產力：生成式 AI 增強員工在報告撰寫、郵件草擬、資訊檢索及程式碼設計等任務之效率，使金融機構得以將更多時間及資源分配至其他重要業務。

(二)競爭壓力

金融機構擔憂在競爭中落後於同業，開始謹慎嘗試 AI 技術（尤其是生成式 AI）。隨著金融機構逐步將 AI 模型應用於客戶端業務，競爭壓力可能加劇。以中長期來看，競爭可促進創新與商業活力，推動商品與服務價格下降，從而提高社會大眾之實際薪資。然而，若

金融機構因競爭壓力而在未充分測試或缺乏安全措施的情況下部署AI，可能帶來潛在風險。

(三)法規遵循

2022 年，英國的金融機構將法規遵循、洗錢防制與打擊資恐（AML/CFT）及客戶審查（KYC）列為 AI 核心應用領域。過去七年，各國對金融機構的監理要求逐步增加，促使金融機構愈加依賴 AI 技術提升其法遵能力，以因應日益複雜的法規環境。

參、 應用案例

由於金融服務業的 AI 應用數據尚不全面且缺乏具代表性之樣本，目前較難對其 AI 應用案例進行深入評估。本報告對現有及新興應用案例進行分類，主要分為產業應用及監理應用。

一、 產業應用

自 2017 年 FSB 報告發布以來，AI 在金融業的應用顯著擴展。然而，隨著監理要求提升，且金融機構對 AI 風險的認識提高，促使其在部署 AI 時採取謹慎態度。許多金融機構傾向於使用較簡單的 AI 模型，旨在輔助而非取代人工決策。

(一)客戶端應用案例

2017 年以來，AI 模型在信用評估、保險定價及行銷等領域應用增加。開放銀行（open banking）等計畫提升數據可用性，使金融機構得以利用 AI 模型評估客戶信用，有助於信用紀錄不足或無信用歷史的客戶，並協助機構預測違約風險。

在行銷領域，AI 可用於強化客製化商品並提高客戶忠誠度。AI 現可協助客戶配對合適的理財顧問，適時為客戶提供所需的商品與服

務。生成式 AI 亦提升數位行銷的效率與規模，能針對特定市場分群生成客製化的宣傳內容。保險業則運用 AI 模型改善保單定價與風險管理，包括透過 AI 來預測個人行為或特定保單風險，並逐步實現保險理賠自動化，但通常仍保留人工審核以確保決策的可靠性。

生成式 AI 提升聊天機器人的互動能力。隨著技術逐漸成熟，聊天機器人的應用亦得以擴展，有望應用於智能理財建議等多種場景。

(二)營運端應用案例

金融機構逐漸於營運端測試並部署 AI 模型，應用範圍包括資本優化、模型風險管理、市場分析及程式碼生成等。相較於客戶導向應用，此類後端應用對提升內部效率及金融體系穩定性具更大影響力。部分金融機構利用 AI 優化市場波動與流動性風險管理。此外，生成式 AI 開啟多項新應用，特別是程式碼生成功能，有助於規模較小機構突破技術人員資源的限制，透過生成式 AI 加速傳統 AI 應用，如詐騙檢測及信用審核。

(三)交易與投資組合管理

量化方法已長期應用於交易與投資組合管理，生成式 AI 進一步擴展應用範圍。部分金融機構利用生成式 AI 分析文本數據，如法說會（earnings calls）[註 14] 會議紀錄等，以評估市場情緒。此外，生成式 AI 工具能結合現有量化工具，實現自動化數據分析與洞察，協助管理者更有效取得並分析相關資訊，提升投資決策的效率與準確性。

(四)法規遵循

自 2017 年 FSB 報告以來，AI 在法律遵循領域的應用顯著增長，應用範圍從反詐騙及打擊洗錢 / 反恐融資（AML/CFT）逐步擴展至調查制裁規避及逃稅行為等。儘管傳統 AI 在識別可疑交易等特定任務上表現優異，但生成式 AI 能整合多種資料來源，自動生成財金犯

罪報告，協助調查人員作業。隨著監理要求不斷提高且金融詐騙事件增加，AI 應用將持續在法遵科技（RegTech）扮演核心角色。

二、 監理應用

金融監理機關正積極運用 AI 技術提升監理效率。調查顯示，監理機關使用 AI 的比例已從 2022 年的 54% 增至 2023 年的 59%。AI 技術亦協助中央銀行執行關鍵任務，例如強化金融機構支付系統的監督、資訊收集以及即時經濟活動分析。部分監理機關使用 AI 分析文本數據（如會議紀錄等）提升監理效率。AI 亦被應用於查核工作中，AI 能從大量文件中提取重要資訊，並協助摘要及撰寫查核報告。

其他應用範疇包括利用生成式 AI 模擬銀行擠兌時，社群媒體的反應及行為模式，以及設計壓力測試情境等。隨著金融機構快速部署 AI，監理機關須緊跟此趨勢並深入掌握其影響，未來 AI 在監理科技中（SupTech）的應用將持續擴展。

肆、 AI 對金融穩定之影響

妥善運用 AI 能為金融機構及其客戶帶來益處。然而，若金融機構迅速引入 AI 而缺乏相應的風險管理與控制措施，加上監理不足及 AI 被惡意行為者利用等問題，均可能對金融穩定構成威脅。本章節將探討 AI 對金融穩定之潛在影響。

一、 監理挑戰

近期 AI 的快速發展可能加劇金融業的脆弱性，包含第三方依賴（third-party dependencies）、市場相關性及網路安全等風險，進而影響金融穩定。儘管部分國家監理機關已掌握金融機構 AI 應用概況，但多為片面數據，且僅涵蓋少數機構，缺乏全面且具代表性的數據，對監理構成重大挑戰。

倘監理機關的 AI 知識及相關技能未能跟上技術發展，監理效能恐受影響。此外，監理機關若未能投入足夠的專業人力及資源，用以全面評估 AI 發展及金融機構 AI 應用之情形，亦可能對金融體系構成風險。一項針對全球 64 個監理機關的調查顯示，多數受訪機關在資訊科學與基本 IT 能力方面仍顯不足。

二、 金融業之脆弱性

AI 相關的脆弱性可能加劇金融市場的系統性風險。首先，金融機構對第三方服務的依賴以及技術供應商市場的高度集中，可能導致國內金融體系與國際供應商之間的連結性加強；主要的 AI 服務供應商僅分布於少數國家，若核心供應商營運中斷或供應鏈受阻，將使金融機構暴露於重大損失風險之中。其次，廣泛使用行為類似或訓練數據來源相同的 AI 模型，可能增加金融市場的相關性，進而加劇市場波動。此外，隨著 AI 應用範圍擴大，惡意行為者得以利用更多攻擊手段，加劇金融機構的網路安全風險。最後，部分 AI 技術的可解釋性有限，且訓練數據品質難以評估，對缺乏健全 AI 治理架構的金融機構而言，可能增加其使用 AI 模型之風險。

(一)第三方依賴與 AI 服務供應商集中度

第三方服務供應商在 AI 供應鏈中，幫助金融機構以較低的時間與成本部署高效 AI。隨著金融業 AI 應用擴大，對 AI 用硬體與基礎設施服務的需求日益增加，加上非專業機構難以負擔訓練 LLMs 的高成本與技術要求，金融業對 AI 服務供應商的依賴加劇，可能為金融機構營運帶來脆弱性。

AI 供應鏈的核心環節，如硬體、基礎設施與數據聚合（data aggregation）[註 15] 等市場高度集中，且部分大型科技公司透過垂直整合，將軟硬體、雲端服務及模型結合為一體。此外，數據聚合市場中，大型數據供應商透過併購小型競爭者擴大控制力，聚合之數據為

金融市場預測模型的重要基礎。由於 AI 服務及相關基礎設施具規模經濟效益，市場集中化難以避免。

LLMs 與生成式 AI 市場中，服務供應商集中化問題日益惡化，主要受以下因素影響：

1.成本與複雜性：訓練 LLMs 需要大量資金及高度專業技術，提高競爭者進入市場的門檻。若終端消費者依賴特定模型，轉換至其他供應商的成本亦將大幅上升。
2.供應鏈限制：AI 所需運算晶片（GPU）持續短缺，導致價格攀升，對資源豐富的企業有利。
3.投資行為：大型科技公司對新創 AI 實驗室大量投資與控制，加深市場集中化。
4.垂直整合：少數企業掌握 AI 供應鏈的多個核心環節，包括軟硬體、雲端服務及模型開發等。
5.通用型模型之需求：市場對通用型模型的需求日益增加，而僅少數大型企業掌握訓練用型模型所需的資源與能力。

然而，以下因素有助於緩解市場集中化，提升市場競爭力：

1.開源模型：開源 LLMs 性能的提升，將降低市場對特定供應商的依賴。
2.技術架構創新：AI 模型訓練技術的突破，能有效降低開發成本並提升效率，有助於市場競爭。
3.GPU 市場競爭：若硬體市場競爭加劇，將有助於增加晶片供應並降低成本。
4.特定用途模型：市場對特定用途 AI 模型的需求，有助於市場多樣性並提高競爭力。

第三方依賴性與 AI 服務供應商集中化對金融市場的系統性影響，取決於 AI 技術的應用範圍。若 AI 成為金融市場的核心角色，

可能削弱金融機構因應網路事件、供應鏈中斷等營運風險之能力，進一步影響金融機構產品及核心服務的穩定性。然而，若金融機構能善加利用特定供應商的專業優勢，提高風險管理效率，亦有助於增強金融體系的整體韌性。

部分金融機構已在核心業務中應用 AI，如信貸決策（credit decisioning）、保險理賠及投資組合管理等領域，並逐步將 AI 應用於風險管理。然而，由於缺乏金融機構 AI 應用之相關數據，目前尚難判斷 AI 於上述應用中的角色與功能。生成式 AI 因操作簡便且科技公司積極推動，正逐步融入金融機構的核心業務。部分系統性重要金融機構已將其應用於價格發現（price discovery）（註 16）、投資建議及風險管理等領域。此外，生成式 AI 有望透過程式碼生成應用，協助開發與維護銀行核心系統。然而，若多數金融機構依賴相同或相似的 AI 模型生成程式碼，可能加劇金融市場的營運脆弱性。

為因應第三方風險，多數監理機關已建立第三方風險管理架構，要求金融機構進行盡職調查（due diligence），持續監控並定期進行風險評估。FSB 出版的第三方風險管理指引（third-party risk management toolkit）亦補充各類評估指標及相關策略。

(二)市場相關性

金融機構可能因使用之 AI 採用相似模型或是數據來源，導致金融市場相關性增加，其原因包括：

1.羊群效應：市場參與者傾向模仿其他機構選擇的數據與模型。

2.網絡外部性：第三方訓練的 AI 模型性能因終端用戶數量增加而提升，推動市場參與者使用相同的第三方模型。

3.選擇有限：僅少數訓練數據及 AI 模型達到可實際應用之水準，供應商集中化限縮選擇範圍。

4.透明度不足：AI 模型供應商未揭露其訓練數據來源，導致終端用戶可能無意間依賴與其他市場參與者相同的數據。

市場相關性可能對金融市場之交易、借貸及保險定價等產生影響。基於相似數據來源訓練的模型可能做出高度相關的預測，即使機構對特定領域進行微調，其核心模型仍相同，從而導致金融機構間的決策趨於一致。若將 AI 廣泛應用於投資組合管理與金融市場之交易，恐在市場下行時加劇波動，增加市場閃崩（flash crash）風險。市場相關性亦可能因風險管理標準趨同而提升，放大市場隨景氣波動的幅度。若 AI 模型風險治理不足或解釋性有限，上述風險恐將進一步惡化。然而，若金融機構主要將 AI 用於提升內部效率，例如文件撰擬、員工培訓及資訊檢索等，對市場相關性的影響則相對有限。

金融市場自動化可能進一步提升市場相關性。隨著自動化交易日益成長，企業逐步採用半自動化的財務管理服務。金融市場亦正積極推動高效、全天候運行的資金流通機制。自動化可提升交易效率，但亦可能放大市場相關性。例如，採用半自動化財務管理之企業，能即時針對相關新聞快速轉移存款等。若多家金融機構在市場壓力事件中同時觸發自動化系統的熄火機制（kill switch）（註 17），可能加劇市場動盪。

另一方面，若將 AI 應用於制定客製化交易與投資策略，或是透過 AI 技術降低客戶進入金融市場的門檻，則可提升金融市場的多樣性，有助於降低市場相關性。資產與財富管理公司早已開始採用客製化的 AI 工具，協助專業投資人決策。

透明度與波動控制機制可緩解 AI 造成的市場相關性風險。例如，監理機關可強化相關透明度規範，市場交易所則可實施熔斷機制（circuit breaker）。上述措施結合完善的監控機制，將有助於防範風險，並於風險發生前採取預防措施。

(三)網路安全

生成式 AI 工具能降低惡意行為者的技術門檻，提高網路攻擊的數量、創新性及成功率，並助長社交工程、商業電子郵件詐騙、惡意軟體開發及身份偽造等惡意行為。另有研究顯示，部分 LLMs 能在無人引導下自動執行網路攻擊。現代 AI 系統大量使用數據的特性，亦進一步擴大網路攻擊的可能性，其形式包括：

1. 數據及模型污染：透過操控 AI 訓練數據或模型參數，影響 AI 產出內容。
2. 指令注入（prompt injection）：攻擊者透過指令（prompt）誘導 AI 洩漏機密資訊。

若惡意行為者針對金融市場基礎設施、中央銀行及系統性重要金融機構等進行攻擊，可能對金融穩定構成嚴重威脅。針對上述挑戰，金融機構與監理機關已開始使用 AI 進行網路異常檢測，並逐步部署生成式 AI 來增強網路防禦，例如改提升事件檢測能力、加速危機因應及例行性任務自動化等，從而將資源投入其他調查工作。同時，生成式亦可用於檢測惡意程式碼，並教育員工了解網路安全標準及相關技術。此外，根據全球網路韌性小組（Global Cyber Resilience Group）調查，多數中央銀行網路安全專家認為，生成式 AI 帶來的安全性益處大於其風險。儘管 AI 的進步將有助於減少網路脆弱性，但短期內，惡意行為者可能從生成式 AI 受益較多，金融機構及監理機關等在使用生成式 AI 時需考量監理要求與風險管理等，進展較為謹慎，而惡意行為者則無需受此限制。

金融穩定委員會（FSB）、國際標準制定機構（SSBs）及各國監理機關已發布多項標準、指引及工具，以協助金融機構管理網路安全風險及危機事件因應。美國財政部近日點出部分金融機構採取的策略供其他金融機構參考，包括：

1.制定 AI 風險管理架構；

2.將 AI 風險管理納入現有的企業風險管理架構（如三道防線（Three lines of defense）（註 18）架構）；

3.在供應商盡職調查（due diligence）中加入 AI 相關議題。

(四)模型風險、數據品質與治理

廣泛應用 AI 可能增加金融體系中的模型風險，尤其在危機期間，模型的驗證、監控與修正更加困難。部分 AI 技術因解釋性不足及問責性挑戰，難以全面評估其應用效果與可靠性。此外，AI 大量使用非結構化數據及透明度不足的訓練數據來源，加上生成式 AI 引發的「幻覺」（hallucinations）（註 19）等難以察覺之新型錯誤，加劇模型管理的難度。

若金融機構未建立健全的治理結構以監控 AI 使用及其數據來源，模型風險及數據品質問題將更難控制。現代 AI 工具的易用性及實用性恐使金融機構在控制機制尚不健全的情況下貿然使用 AI。金融機構有責任承擔 AI 模型風險及數據品質之挑戰，使用 AI 不能免除其遵守相關風險管理要求的義務。

(五)其他脆弱性

本小節將分析近期 AI 發展如何加劇金融詐騙及不實資訊之傳播，並探討 AI 系統若未能遵循法律、監理及倫理架構運行時，可能引發之相關問題。

1.詐騙

AI 助長金融詐騙，導致金融機構因應詐騙的成本不斷增加。生成式 AI 生成語音與影片之能力（如 Deepfake）可能被惡意行為者用來偽造身分證件或個人檔案等。此外，生成式 AI 亦可被用於詐騙金融機構，如偽造保險索賠資料或商業電子郵件詐騙等。

儘管 AI 能幫助金融機構與監理機關打擊詐騙，但短期內，惡意行為者可能更受益於生成式 AI。金融機構與監理機關在應用 AI 時通常採取謹慎態度，而惡意行為者則無此顧忌；另一方面，用於識別 AI 生成內容的檢測技術尚處於初期階段，難以有效因應詐騙風險。除金融機構的努力外，提升大眾的金融及 AI 素養亦是減輕詐騙風險之重要手段。

2.不實謠言

生成式 AI 可能助長高度擬真的謠言傳播，並在極端情況下對金融穩定產生影響，可能引發市場崩跌或銀行擠兌等。2023 年 5 月，AI 生成的美國五角大廈爆炸圖片曾短暫影響股市。

3.校準不良（misalignment）（註 20）

校準不良的 AI 系統可能為實現利潤最大化而進行不符合法律或監理標準的行為。例如，AI 系統可能透過散布不實資訊引發銀行擠兌，同時賣空相關銀行股票以獲取收益。另有研究顯示，AI 可能在未與其他 AI 溝通的情況下合謀推高價格並影響市場。

4.長期影響

AI 的廣泛應用亦可能改變金融體系結構、總體經濟條件及能源使用模式，對金融市場及金融機構產生影響。例如，資源充足的大型金融機構更積極透過 AI 提升內部效率。同時，AI 技術對能源的依賴日益加深，若非使用乾淨能源，將進一步加劇氣候變化風險。目前，上述影響有待進一步研究與理解，但應持續關注。

伍、 結論

2017 年以來，AI 在金融服務業的應用更加普及且多樣。AI 為金融機構帶來效益，包括提升營運效率、加強法遵管理及提供客製化金融產品等。儘管目前以增加營運收入為目標的 AI 應用尚不多見，但隨著技術的快速發

展及 AI 在金融業的深度整合，未來相關機會將逐步增加。同時，監理科技（SupTech）之應用，以及金融機構在防制洗錢及打擊資助恐怖主義（AML/CFT）中之法遵科技（RegTech）應用亦自 2017 年以來顯著增加。

然而，AI 的應用亦可能加劇金融體系中的第三方依賴、市場相關性、網路安全、詐騙及不實資訊傳播等脆弱性，對金融體系產生系統性影響。此外，AI 應用可能對總體經濟、金融市場競爭及能源使用模式帶來長期變化，進而影響金融體系。若上述脆弱性未能有效監控與緩解，恐衝擊金融穩定。監理機關在監控相關脆弱性時，面臨兩大挑戰：（1）AI 技術與應用之快速演進；（2）缺乏金融機構 AI 應用之代表性數據。

AI 發展加速金融體系走向自動化及快速變化的趨勢，突顯監理機關密切監控 AI 相關進展之必要性。現有金融政策架構已涵蓋許多與 AI 應用相關之脆弱性，但仍需進一步完善以確保架構的全面性。目前監理架構要求金融機構管理網路安全及營運風險，並有效因應模型與第三方風險。然而，AI 發展可能帶來超出現有規定之新議題。部分監理機關已針對 AI 制定指引（guidance），以填補現有規定未能涵蓋的議題。此外，未來 AI 技術發展可能帶來更多新興脆弱性與挑戰，顯示監理機關持續監控、研究與政策調整的重要性。

為因應 AI 應用帶來的潛在風險，金融穩定委員會（FSB）、國際標準制定機構（SSBs）及各國監理機關可考慮以下措施：

一、強化資訊透明度

加強 AI 應用及其對金融穩定影響的監控與評估，例如，定期調查金融機構 AI 應用概況、要求公開揭露相關資訊等。監理機關可加強與金融機構、AI 開發者及其他第三方服務提供商的互動，同時借助學術界專業知識，掌握 AI 最新進展。

二、 評估現行監理架構

檢視現行監理架構是否足以因應本報告提及之脆弱性，並分析特定監理架構對公平競爭環境（level-playing field）[註 21] 之影響。

三、 提升監理效率

FSB 可與相關國際標準制定機構協調，促進資訊與實務共享，推動跨國及跨產業合作，增強監理機關的監理能力。同時，監理機關可考慮利用監理科技（SupTech）及法遵科技（RegTech）工具，提升其監理效率。

註釋

註 1： 生成式 AI（Generative AI, GenAI）係一種 AI 系統，能根據輸入指令生成文字、圖像或其他內容，透過學習數據的模式與結構，產生與訓練素材相似但具變化之內容。例如，ChatGPT 即為生成式 AI 的典型應用之一。

註 2： 深度學習（deep learning）係機器學習的分支，以模仿人類大腦的人工神經網路為基礎，能高效提取與分析數據特徵。其應用涵蓋圖像識別、語音辨識及自然語言處理等領域。

註 3： 非結構化數據（unstructured data）係指未以固定格式存儲的數據，如文字、圖片或影片，需要透過 AI 或特殊技術進行分析及處理始能提取有用資訊。

註 4： 嵌入技術（embeddings）係一種將複雜數據轉化為易於電腦處理形式的技術，用於幫助電腦更高效地分析及理解非結構化數據，如文字或影像。

註 5： 變換器架構（transformer architecture）係一種深度學習模型架構，由 Google 於 2017 年提出，透過聚焦於數據中最關鍵的訊息，革新自然語言處理技術，並成為生成式 AI 與大型語言模型的基礎。

註 6： 自然語言處理（Natural Language Processing, NLP）係一種機器學習技術，使電腦得以理解並處理人類語言。

註 7： 大型語言模型（Large Language Models, LLMs）係基於大規模文本數據訓練的深度學習模型，用於實現自然語言的理解、生成及處理。

註 8： 圖形處理器（Graphics Processing Unit, GPU），又稱顯示卡，最初用於圖形渲染的硬體設備，現已被廣泛應用於人工智慧與深度學習領域，能高效處理大規模數據並支持複雜模型的運行。

註 9： 機器學習（Machine Learning, ML）係 AI 的一個分支，透過演算法使電腦從數據中學習並自動改善其性能，常應用於模式識別、預測分析及數據分類等場景，為 AI 技術的核心實現方式之一。

註 10：合成數據（synthetic data）係指利用演算法或模型模擬生成的數據，模擬真實數據的統計特性與結構，用於訓練 AI 模型或進行測試，常用於補充數據不足或保護隱私。

註 11：預訓練 AI 模型（pre-trained AI models）係指 AI 服務供應商已訓練完成的模型，金融機構可直接套用或進行微調以滿足特定需求，無需自行從零訓練模型，大幅降低技術開發的門檻與成本。

註 12：解釋性（explainability）係指 AI 模型之運作邏輯及輸出結果能否被清楚解釋及理解，有助於提升透明度與信任度。

註 13：開源（Open Source）係指軟體或模型的程式碼公開，允許任何人自由檢視、修改並使用，廣泛應用於技術開發。封閉源程式碼（Closed Source）則與之相對，程式碼不公開。

註 14：法說會（earnings calls），全稱「法人說明會」，係企業管理層舉行之電話或線上會議，向投資者說明財務業績、經營狀況及未來展望等。

註 15：數據聚合（data aggregation）係指將來自不同來源的數據彙整、整理並統一格式，以作為後續分析或模型應用之基礎。

註 16：價格發現（price discovery）係指買賣雙方透過交易行為確定商品價格之過程。

註 17：熄火機制（kill switch）係指市場出現極端波動或異常情況時，自動化系統設置之保護措施，用於快速暫停交易或觸發風控流程，以防止損失擴大。然而，若多個系統同步啟動熄火機制，可能引發恐慌，造成市場動盪。

註 18：三道防線（three lines of defense）係由巴塞爾銀行監理委員會（Basel Committee on Banking Supervision, BCBS）提出之企業風險管理架構，包括

第一道防線的業務部門負責之日常風險控制，第二道防線的風險管理及法遵等部門負責監督與指導，及第三道防線的內部稽核部門負責獨立檢查與評估。

註 19：幻覺（hallucinations）係指 AI 生成之內容看似自信且具有說服力，但實際上可能為 AI 自行編造且與事實不符的資訊，容易誤導使用者。

註 20：校準（Alignment）係指 AI 系統的目標及行為是否與人類之法律、倫理及社會價值相一致。校準不良的 AI 系統可能在未充分考量風險與後果的情況下運行。

註 21：公平競爭環境（level-playing field）係指市場參與者在公平條件下進行競爭，避免因規模大小、技術水平或市場地位等而適用不平等規則導致競爭失衡。