澳 門 金 融 管 理 局
AUTORIDADE MONETÁRIA DE MACAU

通函編號 4226/MC011-2026-DSB/AMCM

致：澳門
全體金融機構

收件人參考 V/REF.ª	日期 DATADA DE	寄件人參考 N/REF.ª	日期 DATA 14/10/2026

事　　項
ASSUNTO: **應用人工智能的注意事項**

近年來人工智能（AI）技術發展迅速，生成式人工智能（Gen AI）、AI 代理（AI Agent）等新興技術逐漸為各地金融機構所採用，惟 AI 的應用帶來效益但也伴隨着風險。有鑑於此，本局現透過本通函明確有關應用 AI 技術的注意事項，以協助金融機構完善 AI 管治及風險管理框架，針對 AI 應用實施與風險相稱的管控措施。

一、 公司治理與最終責任

(1) 金融機構在引入 AI 技術前，董事會（對分行而言，指本地高級管理層或負責該分行運作的總行管理層）應明確有關 AI 應用的具體戰略目標及評估其對機構自身的業務影響及價值，並配置適當的專業人員和資源，建立健全的數據治理及 AI 應用管治框架，以及將之納入內部整體風險管理體系，以有效識別、監測、評估及應對有關 AI 應用所涉及的各類風險。

(2) 董事會須對 AI 的應用和有關的風險管理承擔最終責任。為此，董事會須對有關 AI 應用所涉及的風險及其技術局限性有適足的認知，並確保所有參與數據治理、AI 開發、模型訓練、測試、驗證、監控及應用等人員均具備擔任其崗位所需的專業知識與技能，以及持續獲得相應的培訓，以確保相關風險在 AI 技術的快速發展下仍能得到有效的管控。

澳門東望洋斜巷 24 及 26 號
CALÇADA DO GAIO, Nº 24 e 26, MACAU
• 電話TEL : 2856 8288 • 傳真FAX : 2832 5433 • 郵政信箱CAIXA POSTAL : 3017

二、 AI 應用風險分級及管理

(1) 金融機構應根據 AI 應用對客戶或／及機構內部操作流程的影響程度、模型複雜程度，以及業務流程或決策依賴 AI 應用的程度，為 AI 應用進行風險分級。就具備以下特徵的 AI 應用，應界定為高風險應用。AI 決策：

i. 對客戶權益、財產安全構成重大影響；
ii. 對機構業務營運、資訊安全、財產、業務連續性構成重大影響；
iii. 應用於監管合規、客戶盡職調查、反洗錢及反恐融資／反欺詐交易監控、偵測及應對網安攻擊，或對監管合規或關鍵風險管理程序構成重大影響；
iv. 使用大量敏感與私隱數據進行模型訓練、或在部署後將客戶敏感個人資料傳輸至 AI 應用進行運算。

(2) 金融機構應建立完整的 AI 應用清單並確保清單資訊在出現變化時作出及時更新。清單應至少包涵以下資訊：

i. AI 的應用場景及其技術局限性；
ii. 模型類型與演算法架構；
iii. 數據來源、類型與治理機制；
iv. 風險分級結果；
v. 監督與控制機制。

(3) 金融機構應根據有關 AI 應用的風險分級、自動化程度及應用場景，設有適足且有效的人員監督與控制機制，如引入：

i. “人在環中”（human-in-the-loop）機制，在決策過程中人員保留實質控制與否決權，決策方得生效，以防止或減輕 AI 模型作出不當決定而對客戶及機構的影響。此機制尤其應實施於高風險的 AI 應用；
ii. “人在環上／人在指揮”（Human-on-the-loop / Human-in-command），人員監督 AI 系統的運作與輸出結果，並在必要時（如 AI 系統發生異常、決策低於置信度或觸發合規風險條件等情況）及時介入、修正或啟動緊急停用機制。

三、 AI 應用的全生命週期管理

(1) 金融機構應將健全的數據治理作為部署 AI 應用的前提。機構應貫徹風險為本原則，建立涵蓋模型研發、數據預處理、模型訓練、模型可解釋性與穩定性驗證與評審、部署與運行監控，以及終止後的數據存儲與銷毀等環節的 AI 系統全生命週期管理機制。此外，金融機構應設有紀錄保存、變更管理審查及應急措施，並因應技術迭代、業務轉型及外部環境變化，定期評估及適當修訂有關生命週期的管控措施，以確保措施的有效性及前瞻性。

(2) 針對模型研發數據預處理環節，金融機構應做好有關數據清洗、數據標註、數據格式統一、敏感資訊（如客戶個人資料或機構內部敏感文件）梳理及數據脫敏等各類數據加工操作。機構應確保所有用於訓練、驗證或測試 AI 模型的數據來源合規、內容準確，以及有關的預處理流程可追溯。

(3) AI 模型的可解釋性應與其對業務的重要性及風險水平相稱。機構應可清晰闡明、驗證及記錄 AI 模型的運作機理、其預測或決策過程的核心邏輯。

(4) 金融機構應確保 AI 模型決策在相同條件下，具備合理的重現性與穩定性，其輸出結果應保持高度一致或在可接受的合理區間內。若模型的重現性或穩定性未達預期，機構應採取相應的風險緩釋措施，包括但不限於：適當降低模型複雜度、剔除缺乏業務合理性的特徵變量，或在關鍵決策節點實施更嚴格的人員監督與控制機制。

(5) 高風險 AI 應用部署前，應由獨立於模型開發與訓練的團隊（或第三方）進行驗證，驗證範圍應至少涵蓋概念合理性、數據質量、模型的安全、可靠及技術局限性、性能表現，並向董事會報告有關驗證結果及所設置的風險緩釋措施的有效性。對於非高風險的 AI 應用，驗證範圍可按應用的業務需求作適當調整及由具備相關資歷的內部人員進行。

(6) AI 模型決策質量高度依賴數據、業務情境與演算法架構的穩定性。金融機構應建立健全的 AI 模型動態監控機制，持續追蹤在模型部

署後的整體運行情況及環境的變化，一旦監控指標超出預設的效能閾值或出現重大預測偏差，又或發現以下情況，機構應立即啟動模型風險緩釋與補救程序，及時重新進行模型校準、訓練或開發。

i. “數據漂移”或數據結構性異常。生產環境輸入的數據的統計特徵及統計分佈、底層數據源結構（數據的組織架構、欄位定義或資料型態、擷取來源等）、數據質量（異常高的缺失率、噪聲與格式錯誤等），與訓練階段原始數據相較發生偏離或實質性變化；
ii. “概念漂移”。現實世界中的客戶行為或市場環境發生轉變，輸入的數據與預測的目標結果之間的統計映射關係發生了根本變化，原有的關聯性不再適用；
iii. 環境結構性差異。總體經濟環境、市場機制、消費行為模式或監管法規與原始參考基準相較出現重大脫節；
iv. 業務範圍擴展。將 AI 應用擴展至新的業務模式、客群範圍、新市場或新產品線，導致實際應用場景超出原模型設計的預期邊界；
v. 模型本質重大變化。模型核心算法、推理引擎或訓練框架進行了重大版本升級、演算法底層架構變更，或影響模型決策權重的核心參數進行了調整。

(7) 為防範 AI 應用因高度自動化與高速運算引發風險傳導與連鎖反應（Cascade effect），金融機構應建立健全的“失效安全機制”（Fail-Safe Mechanisms）與業務連續性計劃。針對高風險 AI 應用，機構應設定明確的操作邊界與風險容忍度閾值。一旦系統運行超出閾值或發生重大異常，須立即啓動“緊急中斷開關”（Kill-Switch）或人工干預程序，並確保具備將原有借助 AI 應用執行的業務，快速切換至備用系統或人工替代流程的能力，以將營運或服務中斷的影響降至最低。此外，機構並應定期就上述機制與計劃進行演練。

(8) 為降低因 AI 系統終止運作後可能衍生的個資外洩、私隱侵害及技術資產殘留等潛在風險，金融機構應就終止 AI 的應用，建立完善的數據清理與退場機制。除須遵守有關數據保存規定、及為滿足審計與爭議處理所需的必要數據外，機構應及時對用於該 AI 模型的訓練、驗證的相關數據、模型權重、特徵向量或暫存檔等進行安全

銷毀、清理或去識別化，確保資料無法被還原。前述的程序如涉及第三方服務供應商，應確保其妥善落實清理，並應留存銷毀軌跡或證明以供機構查核。

四、 第三方服務供應商的風險管理

(1) 金融機構引入第三方 AI 服務前，應明確界定與第三方之間的風險管理及監控責任分工，惟機構董事會仍須對 AI 應用產出的結果與潛在風險承擔最終責任。機構除須按現行《外判管理指引》的相關監管要求外，亦應針對 AI 應用的特有技術性質和風險，評估服務供應商的 AI 及數據治理機制，包括模型開發、訓練和驗證標準，以及訓練數據的來源、合規性與偏見防範措施等。

(2) 金融機構應確保高風險 AI 應用於部署前或作出重大變更前，已由獨立於業務及對接 AI 服務供應商的內部人員，或由獨立於供應商的外部機構（兩者均須具備相關技術安全評估資歷），進行有關應用的代碼審查、漏洞掃描及安全測試，並應定期排查及防範惡意組件注入等安全威脅。

(3) 金融機構應識別並管理因依賴單一或少數 AI 服務供應商而引致的集中度風險，並建立相應的風險緩釋措施及營運韌性計劃。同時，機構應落實嚴謹的資料處理與安全防護措施，透過合約明文禁止服務供應商在未經機構事前書面授權下留存數據，或將之用於其自身模型的訓練或再利用於其他用途，以防止敏感資訊不當外洩。

五、 面向客戶的保障與透明度

(1) 金融機構應就 AI 應用向客戶提供適足的透明度及保障。在向客戶提供服務前，應主動明確告知客戶該服務涉及 AI 的應用，並以淺顯易懂的語言向客戶說明 AI 在服務中的用途、目的、基本決策邏輯與限制。針對面向客戶的高風險應用，機構應向客戶提供替代方案（如引入“人在環中”或純人工替代流程）。另外，機構應設有暢通的申訴渠道與爭議處理機制，確保客戶能對 AI 決策提出異議，並獲得及時、公正的人工覆核與處理。

(2) 金融機構應落實相應機制以避免 AI 決策產生歧視或偏見。機構應於 AI 模型開發環節及部署上線後，定期進行偏見與公平性評估，審查模型決策有否對特定群體構成不公平。一旦識別偏差，應即時糾正。若於模型訓練及決策過程直接、或經由代理數據間接衍生利用（指模型雖移除了性別、種族、年齡等敏感標籤，卻利用與前述高度連動的代理數據在底層演算法中重新拼湊出特定人群的特徵輪廓）易引起歧視的個人特徵（如種族、性別、年齡、宗教等敏感個人數據），應具備客觀、實質且非歧視性的商業合理原因，且通過專項法律與合規審查。機構並應定期進行模型驗證以確保模型決策符合公平性指標。

最後，為確保金融機構的 AI 應用已實施與其風險水平相稱的風險管控措施，機構擬落實高風險 AI 應用前，應先與本局作充分溝通。如有疑問，請派員聯絡本局銀行監察廳蔡法中先生或曾家豪先生（電話：8395 2357／8395 2312，電郵:dsb@amcm.gov.mo）。

耑此。順頌

商祺

澳門金融管理局

行政管理委員會委員

劉杏娟

銀行監察廳副總監

何仲華