

(翻譯本)

本局檔號：B1/15C

B9/29C

致：所有認可機構
行政總裁

敬啟者：

應用人工智能的高層次原則

香港金融管理局（金管局）於2019年第3季調查銀行使用人工智能的情況，發現不少銀行正在或計劃應用人工智能，其使用範圍亦由面向客戶的服務（例如：聊天機械人及個人化市場推廣）延伸至內部運作及風險管理範疇（例如：自動化操作、網絡及詐騙風險管理）。調查結果的詳情將會很快公布。

人工智能的應用逐漸普及，既為銀行創造機會，亦會帶來風險管理的新挑戰。因此，金管局認為現時適合為銀行業提供應用人工智能的指引。本通告所載的高層次原則，是在參考業界的穩健做法及主要海外監管機構的相近原則後訂立，銀行在設計及應用人工智能及大數據分析時應予考慮採納。金管局明白過份規限或硬性的規定，或會窒礙人工智能相關科技進一步發展，因此金管局將這些原則制定為高層次原則。銀行可按情況相稱地採納這些原則，以反映其人工智能應用程式的性質及所涉風險水平。

管治

1. 董事局及高級管理層須為應用人工智能所引致的結果負責——部分人工智能應用程式具備從經驗及例子中自學的能力（例如：透過強化學習及深度學習），並可自動代銀行作出決定。銀行董事局及高級管理層應明白自身仍須就所有人工智能應用程式作出的決定負責，因此應確保銀行具有妥善的管治框架及風險管理措施，以監察其應用人工智能的情況。銀行亦應清楚訂明有關開發及監察應用人工智能運作時三道防線的角色與責任。

應用程式的設計與開發

2. 具備足夠的專業知識——由於人工智能應用程式的設計及開發牽涉特定的專業知識，因此銀行應確保相關開發人員具備所需的專業能力及經驗。高級

管理層應確保訂立有效機制監督有關人員，亦應制定適當計劃招聘、培訓及留住具有合適技術的員工。

3. 確保人工智能應用程式有恰當程度的可解釋性——穩妥可靠的人工智能應用程式對所有相關方面都應該是可解釋的（即不能以「黑盒」作藉口）。因此銀行應於設計階段採取足夠措施，確保應用程式具適當程度的可解釋性，與有關程式的重要性相稱。
4. 採用高質量數據——人工智能應用程式的準確性及表現很大程度取決於用作訓練人工智能模型的數據，因此銀行應制定有效的數據管治框架，確保使用高質量和相關的數據。例如，銀行可根據適當界定的數據質素標準（包括數據準確性、完整性、適時性及一致性等因素）評估有關數據的質量，經評估識別出的數據質量問題應匯報至相關人員予以及時糾正。
5. 嚴格核實模型——銀行應嚴格核實及測試已受訓的人工智能模型，確保準確合適才投入運作使用。此外，銀行亦宜安排獨立方（例如：第二或第三道防線，或外部顧問）參與模型核實的過程。
6. 確保人工智能應用程式的可審計性——人工智能應用程式的使用結果應可以接受持續稽查，一旦出現事故或不理想情況時，亦可按需要收集證據協助調查。因此，銀行應於設計階段建立足夠的稽核紀錄及相關文件紀錄，而有關稽核紀錄及文件應保留一段合適時間，以確保可予稽查審視。
7. 對第三方供應商實施有效的管理監察——若銀行倚賴第三方供應商開發人工智能應用程式，應按照本通告所載適用原則對其進行適當的盡職審查，銀行亦應採取有效的供應商管控措施，包括定期檢視供應商的服務以管理相關風險。
8. 秉持道德操守、公正及具透明度——銀行應確保人工智能所作出的決定不會歧視或無意間對某客戶群造成偏見。人工智能應用程式的使用須遵守其企業價值和道德操守準則，以及堅守保障消費者的原則。為提供透明度及增加消費者使用人工智能服務的信心，銀行應在提供相關服務前，向消費者披露該服務是以人工智能技術推動和當中所涉及的風險。

持續監察及維護

9. 定期檢視及持續監察——由於人工智能應用程式能從生產數據中學習，其模

型行為於投入運作後或會有變化，因此銀行應定期檢視（例如：在適當情況下重新核實有關人工智能模型）及持續監察，確保應用程序繼續按照原定設計運作。

10. 遵守保護數據規定——鑑於人工智能應用程式需使用大量數據的特性，銀行應採取有效的措施保護有關數據。若人工智能應用程式需收集及處理個人資料，銀行應遵守《個人資料（私隱）條例》及其他適用的本地與海外監管規定。此外，銀行應在適當的情況下，使用已剔除敏感資料的數據，而非可識別個人身分的資料。
11. 實施有效的網絡安全措施——使用人工智能應用程式或會構成新的網絡安全風險，例如數據中毒攻擊及對抗性攻擊，能透過操縱數據以影響人工智能模型。銀行應確保網絡安全措施能有效應對這些攻擊，亦應時刻對新出現的網絡風險保持警惕，並掌握相應防護措施。
12. 風險緩減措施及應變計劃——即使最穩健的人工智能應用程式仍有可能產生非預計的後果。銀行除了實施適當的風險緩減措施（例如：加入人手流程的機制、設定審慎的風險限額及進行質量保證抽查），亦應制定應變計劃，以能迅速暫停人工智能應用程式運作及啟動後備程序（例如：採用人手干預或傳統流程）。

考慮到針對人工智能使用的國際監管標準及業界發展迅速，金管局會定時檢視上述原則，並適時向銀行提供進一步指引。針對應用人工智能時涉及的消費者保障事宜，金管局計劃另行發出指引。

若對本通告有任何疑問，請聯絡趙紫瑋先生（2878-1389）或鍾瑋羚女士（2878-8533）。

助理總裁（銀行監理）
陳景宏

2019 年 11 月 1 日