人工智慧在金融詐騙中的濫用風險與防範策略

資訊處

許芷菱

中華民國 114 年 6 月

作者任職於中央銀行資訊處，本文內容純屬個人意見，與服務單位無關，如有錯誤概由作者負責。

摘要

隨著人工智慧技術在金融業的廣泛應用，詐騙手法亦因其演進而變得更加隱蔽、精準與難以辨識。生成式 AI、深偽技術（Deepfake）與自動社交工程等工具已成為不法分子從事金融詐騙的新利器。這些 AI 技術不僅提升詐騙效率，更增加監管機關在即時偵測與風險控管方面的挑戰。儘管目前已有多種防範機制與監管措施應運而生，包括多重身分驗證、行為分析與異常偵測等，但仍因法規滯後、跨境執法困難與 AI 模型的黑箱特性等因素，使監管工作在實務上面臨諸多障礙。本文將探討 AI 在金融詐騙中的應用風險、現行防範策略、監管挑戰，並提出因應對策，以期在促進 AI 發展的同時，維護金融市場的安全與穩定。

內容目錄

圖表目錄

壹、 緒論

在當前的數位時代，人工智慧（AI）技術迅速發展並廣泛應用於金融領域，從客群經營、智能風險控管、流程的自動化與數位化、到大數據分析，每一環節都離不開人工智慧。金融機構利用 AI 來提升業務效率、降低營運成本，並優化服務的品質。然而，人工智慧技術的普及也帶來了許多潛在的風險，特別在金融詐騙領域。人工智慧技術在金融詐騙中被濫用的風險與日俱增，人工智慧的進步使得詐騙手段更加複雜和難以偵測，對金融市場以及個人投資者造成了巨大的威脅。

金融詐騙，自古以來便是一個持續存在的問題，而隨著 AI 的發展，詐騙手段也越加的多樣化。傳統的詐騙依賴人工的操作，詐騙者需耗時耗力的進行詐騙活動。而現今，詐騙者可以利用 AI 的技術，像是利用機器學習、自然語言處理等技術，創造出難以辨識的詐騙行為。舉例來說，詐騙者可以利用伴隨 AI 技術出現的深偽技術 (Deepfake) (刑事警察局公共關係室, 2024)，假冒知名投資人進行詐騙活動。因為 AI 可以模擬人類的行為與語言，讓各種詐騙更加的真實與具有隱蔽性，增加了受害者識別真偽的難度，也使得金融機構以及監管部門都面臨了前所未有的挑戰。

AI 在金融詐騙中的濫用風險可以從多方面著手，除了先前提到的深偽技術，詐騙者也可以利用生成式 AI 生成虛假的金融商品，甚至是操控市場數據，創造虛假的市場信號，誘導民眾上鉤。這些 AI 詐騙活動不僅影響了投資者的財產安全，甚至會威脅金融市場的穩定性。在高度數位化的金融環境中，AI 的濫用提高了詐騙者的隱蔽性，令傳統的反詐騙措施難以有效的應對。

為了有效防範人工智慧技術在金融詐騙中的濫用，金融機構需要採取多層次的防範策略。這些策略包括加強數據的安全管理、建立完善的身分驗證機制以及採用先進的人工智慧技術來識別和預防潛在的詐騙行為。此外，不僅需要在技術層面提升防範能力，還應從監管、法律、教育等多方面進行布局。要確保 AI 技術的正當應用，並加強對各種新型詐騙行為的監控與預警能力，同時，金融機構還應培養員工的風險意識，確保員工在遇到不論是客戶遇到詐騙行為或是自身落入詐騙的陷阱中時能及時發現並做出正確的應對措施。金融機構、監管機構、法治機構以及其他利益相關者必須通力合作，制定出一套有效的防治策略。

本文將探討 AI 在金融領域中所扮演的角色以及在金融詐騙中的濫用風險與防範策略。首先，我們將先了解 AI 在金融行業中的應用；

接著探討 AI 如何被不法分子用於金融詐騙，之後探討金融機構與監管機構在面對這些新型詐騙遇到的挑戰；最後，提出防範策略，幫助金融機構在面對 AI 詐騙時能有效的應對，確保金融市場的穩定與安全。本文旨在透過這些分析與建議，讓金融產業在 AI 的技術應用與風險控管之間找到平衡，促進金融領域的穩定與健全發展。

貳、 人工智慧在金融行業中的正面應用

根據金管會於 2024 年 5 月的調查報告指出，我國目前在 377 家金融業者中，有 108 家導入人工智慧(AI)應用 (財團法人金融研訓院, 2024)，其中銀行業比例高達 74%，而證券期貨業雖然只有 18%，但有導入生成式 AI 的比例卻占 37%。這顯示出人工智慧(AI)在金融業的應用已經相當廣泛且比例高。在金融行業中，人工智慧 (AI) 透過深入分析大量數據，挖掘出顧客行為模式、解析變幻莫測的市場局勢，並輔助決策過程。目前，AI 已廣泛應用於風險控管、客製化的金融服務、智能投資顧問，甚至是詐騙檢測等領域。特別是在風險管理的部分，AI 透過快速地分析巨量資料，協助金融機構識別潛在風險並即時進行調整。此外，AI 技術還能預測市場走向，幫助投資者做出更客觀精確的決策，推動金融業的智能化發展。這

些應用不僅提升了效率，也有助於優化決策流程，從而提供更好的金融服務。接下來，我們將進一步探討人工智慧(AI)在金融領域中的具體應用，涵蓋以下幾個重要領域：

一、 風險控管

根據 SAS 調查指出，人工智慧(AI)最大中應用於預測(54%)和最佳化(51%)，其次是，其次是「機器學習」(34%)、「機器人流程自動化」(29%)等技術。 (美商賽仕電腦軟體(SAS), 2019)。透過分析大量數據，AI 可以協助銀行在面對複雜的市場動態和多樣化風險時，識別潛在風險並提供有效的風險控制措施。例如，AI 可以從大量的信用資料和市場波動數據中，預測信貸風險、詐騙行為及市場變動。此外，機器學習演算法的應用，能提升風險預測的準確性，減少人為錯誤，進而穩定決策過程並提高準確性。對於金融業而言，AI 不僅能通過分析客戶的交易行為識別欺詐行為，還能顯著降低金融機構所面臨的風險。

二、 個性化金融服務(智能投資顧問、智能客服)

人工智慧（AI）在個性化金融服務領域，特別是投資顧問和客戶服務方面，正發揮日益重要的作用。機器人理財[1]有很多類型，其一是利用大量蒐集的資料，根據演算法進行風險管控、市場報價、投資效益等分析，回報結果供投資人參考；而有些類型則是依據投資人的投報風險接受度、財務狀況、投資期限、帳戶類型、交易成本等資訊分析出客戶合適的投資組合 (陳安斌;陳莉貞;蘇秀玲;郭怡君 , 2018)。這種計算自動化財富管理或投資建議提供的服務，跟傳統的理專相比，自動化、數據化的投資建議，降低了人為介入的影響，同時也提升了服務效率並降低成本。此外，人工智慧也可以實現自動化交易，從而提高交易的效率和準確性。通過分析市場數據和趨勢，AI 可以自動執行交易指令，並根據市場變化進行調整，這不僅降低了人工操作的風險，還提高了交易的速度和準確性。

而智能客服的提供，不僅提供 24 小時的諮詢服務，還包括常用服務圖卡提醒等服務，這不僅提高了客戶的滿意度，還降低了金融機構的運營成本。智能客服能夠快速解答客戶的常見問題，節省了客戶的時間。此外，當民眾需要變更資料時，智能客服使這一過程

[1] 理財機器人是什麼？一篇比較智能理財優缺點與分析 3 大功能
https://bank.sinopac.com/sinopacBT/webevents/ibrain/blogs_detail_022.html

更加方便，無需等待上班時間即可完成操作。下圖是玉山銀行智能客服的操作頁面。

圖表 1 玉山銀行智能客服[2]

三、 銀行後臺業務自動化與系統更新

生成式 AI 結合深度學習技術，應用於會計審計領域，能在會計任務中實現自動化，協助稽核部門自動執行財務審核工作，極大提高了工作效率，減少了人工錯誤。其次，對於系統更新的部分，許多依賴於傳統程式語言（如 COBOL）的銀行核心系統，或者日益更新的程式語言，生成式 AI 能夠靈活地輔助撰寫，幫助銀行克服人才

[2] 玉山銀行智能客服 https://robot.esunbank.com.tw/

短缺的問題，加快系統維護和升級速度。這不僅降低了技術成本，還能改善銀行的整體運營效率[3]。

總結來說，人工智慧（AI）在金融業的應用正迅速增長，並且在多個領域發揮著關鍵作用。無論是在風險管理中利用大量數據預測和控制潛在風險，還是在個性化金融服務中提供智能投資顧問和自動化交易，AI 都能顯著提升效率和準確性，並優化客戶體驗。同時，生成式 AI 在銀行後臺業務的應用，如會計自動化和舊有系統的更新維護，進一步減少了人為錯誤、提高了工作效率並降低了成本。隨著 AI 技術不斷應用於金融業，對於銀行本身或是客戶，都將帶來巨大的效益。

然而，儘管人工智慧(AI)帶來了眾多優勢，隨之而來的風險與挑戰也不可忽視。尤其是在金融詐騙領域，AI 技術的濫用可能會引發一系列安全隱患，對金融機構和消費者帶來潛在威脅。接下來，我們將探討人工智慧在金融詐騙中的濫用風險，以及如何應對這些挑戰。

[3] 生成式人工智慧對金融服務的價值核心
https://www.ey.com/zh_tw/insights/ai/core-values-in-generative-ai-for-financial-services

參、 人工智慧在金融詐騙中的濫用風險

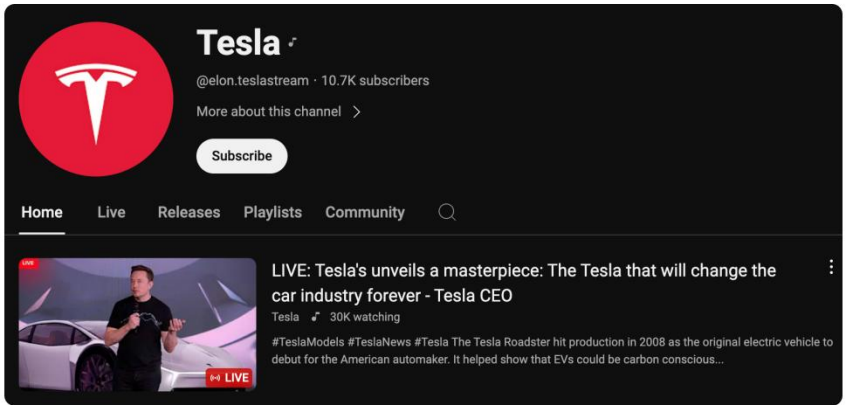
隨著人工智慧的快速發展，詐騙者開始運用機器學習、深度學習及生成式 AI 等技術來構建更加複雜的詐騙手法。機器學習被用來分析大量的金融交易資料，識別出具有高潛力的目標群體。詐騙者透過分析消費者的行為模式，創建出精確的假冒攻擊方案，能夠精確模擬合法金融交易或行為，讓受害者無法輕易辨識。

在金融詐騙領域，詐騙者也有可能利用 AI 強大的功能，進行難以辨識的詐騙手法，從而對金融機構及客戶造成重大損失。以下分別以實際案例探討詐騙者如何利用人工智慧技術進行詐騙、以及這些行為到底會造成何種風險。

一、 偽造身份與虛假資料生成

生成式 AI 在金融詐騙中被廣泛濫用於生成假訊息與偽造身分上。詐騙者一方面利用 AI 生成看似真實的假訊息，並將其廣泛應用於電子郵件或社交媒體詐騙中，誘使目標點擊釣魚連結或提供敏感個人資料。另一方面，AI 技術也被用於建立虛假的身分識別系統，犯罪分子透過深度偽造技術，製作出與真實身份幾乎無異的文件或證件，進而獲取金融服務的授權或操作權限。

有一個實際案例來自網站 OnlyFake[4]，該網站利用神經網絡技術，僅需 450 元台幣（約 15 美元）就能生成逼真的假身分證。用戶可以在這個平台上創建各種假的身分證明文件，包括駕照，並自訂姓名、地址及簽名等個人資料。這些 AI 生成的假身分證非常真實，足以突破許多身份驗證系統的檢測，甚至包括一些加密貨幣交易所（如 OKX）和銀行。這樣的技術規模和可獲得性讓詐騙活動變得更加容易，幾乎任何人都能在短短幾分鐘內生成假身分證，從而繞過許多線上驗證系統，創造許多虛假的銀行帳戶，對整個金融業造成非常大的損失。



圖表 2 馬斯克的影音檔[5].

[4] 暗黑 AI 另一「妙用」：黑網 OnlyFake 用神經網路，450 元就能生成逼真身分證 https://www.inside.com.tw/article/34129-onlyfake-fake-id

[5] "Deepfakes of Elon Musk Are Pushing Crypto Giveaway Scams on YouTube Live." https://www.engadget.com/deepfakes-of-elon-musk-are-pushing-crypto-giveaway-scams-on-youtube-live-200700886.html

2022 年，駭客利用深偽技術(DeepFake)偽造馬斯克等知名人物的影音檔(圖 2)，推銷虛假的加密貨幣交易網站，詐騙投資人的資金。圖 2 瑣事的影片即是一例，該影片看似在馬斯克在特斯拉舉辦的活動中發表演說。影片中，DeepFake 生成的馬斯克點選某個網址，將虛擬貨幣儲存至指定的電子錢包，並聲稱用戶會獲得雙倍的加密貨幣作為回饋。影片中的馬斯克形象、聲音、動作可以假亂真，顯示深偽技術已成為詐騙者操控信任與製造假象的重要手段[6]。

因疫情的影響，金融科技（Fintech）迅速推動了生物識別技術的應用，特別是在遠距辦理服務中，取代了傳統的臨櫃操作。這些生物識別技術，如人臉識別、指紋掃描和聲音識別，讓用戶能夠無需親自前往銀行或金融機構，即可完成身份驗證和金融交易 (謝昀澤, 2022)。然而，隨著深偽技術(DeepFake)的興起，詐騙者開始利用這些生物識別系統的漏洞進行攻擊。利用 AI 生成的假人臉或偽造的語音，犯罪分子能夠成功繞過生物識別驗證，冒充真實用戶進行詐騙。這種情況使得金融業面臨更大的挑戰，因為傳統的生物識別技術可能無法有效防範這些精密的深偽技術(DeepFake)攻擊。因此，

[6] YouTube 平臺出現馬斯克 Deepfake 影片企圖詐騙加密貨幣，吸引 3 萬人同時收看 | iThome

金融機構必須不斷升級安全系統，並結合多重身份驗證措施，以應對深偽技術(DeepFake)所帶來的安全風險。

二、 社交工程詐騙 (Social Engineering)

當前，社交工程詐騙已成為金融詐騙中的一大威脅，尤其是隨著人工智慧技術的進步，這些攻擊變得更加複雜和難以防範。駭客利用 AI 技術來自動化和優化釣魚郵件、利用深偽技術來假冒他人，使得詐騙信息更加真實和具有說服力。舉例來說，一家軟體公司的員工成為自稱是該公司 IT 團隊成員的人的誘導員工點擊看似合法的連結來解決薪資相關問題。員工點擊了該連結，並被引導到偽造的登錄頁面。然後，未經授權的個人又結合 AI 假造實際 IT 工作者的聲音，以獲取獲得登陸頁面所需的多元身份驗證。未經授權的詐騙者成功進入員工帳戶並竊取了 1500 萬美元 (Marsh & McLennan Agency LLC, 2025)。AI 讓駭客在傳送惡意郵件和訊息時沒有了語言翻譯問題，同時也能避免文法和用字錯誤，讓偽造的內容更加逼真，此外，也可以更加聰明地繞過傳統的安全驗證，這不僅提高了詐騙的成功率，還使得攻擊規模更大、速度更快，對金融機構和個

人造成了嚴重的安全威脅 (Santhana, 2025)。因此，了解和預先防範這些 AI 加成的社交工程詐騙對於保護金融系統的安全至關重要。

下表為 AI 加強型的社交工程詐騙的示意圖

	要素	說明
資訊來源	社群資訊蒐集	透過網路爬蟲、社群平台或駭客手法，大量蒐集目標個資、聲音、影像及社交關係，用以建立詐騙模型的基礎資料
技術應用	Deepfake / 深度學習	仿冒語音、臉部影像或虛構訊息，生成極具欺騙性的詐騙內容(郵件、影音、虛假連結)
攻擊手法	假客服 / 騙點連結 / 偽登錄頁面/假影片	結合多種誘導手段

圖表 3. AI 加強型的社交工程詐騙的示意圖

三、 AI 幻覺

隨著人工智慧(AI)的快速發展，用 AI 生成程式碼已更加普遍，對於公司加速開發以及提供更多元的創新力有巨大的貢獻。根據研究機構 Gartner 表示，約有 60%的企業在採用 AI 來開發程式碼，而

其中有 35%的企業更應用於核心產品的開發[7]。AI 自動生成程式碼片段，讓開發者能更加專注於更高層次且難度高的問題解決。

然而，利用 AI 生成程式碼也有其風險。由於訓練模型用不同來源的資料進訓練，所以其生成的程式碼可能有潛藏的安全漏洞或是使用不存在的模組或是函式庫進行撰寫，此稱之為套件幻覺 (Package Hallucination) (Charlotte Thompson&Tiana Kelly, 2023)。 而當駭客發現存在套件幻覺時，便可以創建與幻覺程式庫相同名稱的惡意函式庫，誘導開發者使用此惡意套件，讓系統不知不覺被惡意程式注入，進而對系統造成安全危害。

四、 詐騙客服與詐騙機器人

隨著人工智慧技術的迅速發展，尤其是在深度學習領域，詐騙手法也日益進化。Deepfake 技術，最初用於娛樂產業中合成虛擬影像，現在已被不法分子應用於詐騙活動中，尤其是在模擬客服人員和機器人方面。詐騙者透過這項技術製作出高度仿真的影片跟語

[7] AI 寫程式碼恐隱含威脅　審慎避免軟體供應鏈風險
https://www.netadmin.com.tw/netadmin/zh-tw/viewpoint/6A995C7ED3694834BC7A72B93E28CA5A

音，使其看起來像是真正的客服代表或專業人士，並進行詐騙行為。

在詐騙過程中，Deepfake 可以生成偽造的客服影片，這些影片中的人物會進行高度模擬真人，甚至能精確模仿特定人的口音、語調或肢體語言。這使得受害者不懷疑其真實性，認為他們正在與真人進行交流。而在客人認為其真實性的情況下，不自覺地洩漏了敏感的資訊或是進行資金轉帳的動作，從而對金融機構以及客戶造成損失。

詐騙機器人結合自然語言處理技術，能夠進行更佳的擬真對話。與傳統的機器人相比，結合 Deepfake 的詐騙機器人能夠更自然、更具說服力地與受害者互動。這種情境不僅讓受害者難以識別詐騙的本質，還可能因為影片和語音的高度仿真性，讓詐騙者的行為更難以被檢測與追蹤。

根據 CNN 在 2024 年 2 月 4 日報導的文章[8]，在香港爆發了一起假冒財務長的深度偽造詐騙案件。一家跨國公司的財務人員被

[8] Finance worker pays out $25 million after video call with deepfake 'chief financial officer' https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html

騙去參加一場視訊會議，視訊會議上的每人看起來都非常真實，這讓原本收到來自首席財務長的訊息感到懷疑的員工最終相信了這封釣魚訊息。這名員工最終同意匯出了約 2560 萬美元的資金。

詐騙客服和詐騙機器人之所以會成功，正是其利用了 Deepfake 技術突破了以往人工客服或機器人無法提供的真實感和情感互動。未來，隨著技術的進步，這類詐騙行為可能會更具隱蔽性與危險性，因此在防範詐騙的同時，監管單位也要即時更新更多元的詐騙手法，供民眾去了解並提前預防。

五、 AI 優化的黑市交易

AI 對於交易市場無疑是一把雙面刃：它讓合法交易能夠更高效、精確地進行預測和操作，但同時也使非法交易愈發猖獗，尤其是在黑市這類監管薄弱的領域。自動化分析大量數據，提供準確的市場價格並自動匹配交易對象，從而提升效率；同時，透過精確的行為分析與數據加密，減少了在交易過程中被監控或被發現的風險，增強隱蔽性；除此之外，AI 更能幫助犯罪份子預測並識別執法

部門的監控行為，及時調整交易策略，有效控管風險。這些能力使得黑市交易變得更加高效率且難以被監控。

根據 Chainalysis[9] 所發布的報告，加密幣詐騙在 2024 年達到新高，預計總額將達到 99 億美元，並有可能進一步上升。報告指出，所謂的殺豬盤(Pig Butchering)詐騙，這是一種結合浪漫詐騙和投資詐騙的手法，詐騙者透過建立更具說服力的虛假人設、自動化訊息發送，或分析受害者的心理弱點等手段，進一步提高詐騙的精準度與滲透力，使受害者更難辨識詐騙的本質。截至 2024 年底 ，已呈現出接近 40% 的增長，詐騙者藉由社群媒體應用建立關係，然後以假冒的投資機會欺騙受害人。

在某些高頻交易或金融投資平台，利用 AI 建立虛假的交易指令或分析數據，從而進行詐騙。舉例來說，詐騙者利用機器學習演算法分析金融市場的走勢，生成一組假訊息來操控投資者進行不利的交易，從中牟取不正當利益，最終獲得大量的資金 (Butts, 2025)。這些行為表明，加密貨幣已經成為黑市交易中一個至關重要的工具，尤其是在詐騙活動中。因為加密貨幣結合 AI 技術，讓交易無法

[9] Chainalysis 是一家區塊鏈數據平台公司，為公司提供數據分析、追蹤工具和區塊鏈研究等相關服務。

被輕易破解。與此同時，AI 又能協助黑市詐騙交易者透過數據分析，發現市場的需求提前布局。AI 技術可以說是打破國界，讓跨國的詐騙行為更加容易且具隱蔽性，這也讓國際間要合作打擊這些黑市交易更加困難。

人工智慧在金融詐騙中的濫用對金融市場和監管體系帶來了前所未有的挑戰。首先，AI 技術使得詐騙行為更加隱蔽和高效，傳統的監控和反詐騙手段往往無法有效識別 AI 所生成的虛假信息或身份。詐騙者利用 AI 進行的精確攻擊使得辨識和防範變得愈加困難，金融機構在風險防範方面面臨更高的挑戰。其次，AI 濫用對投資者的安全構成嚴重威脅。無論是虛假交易還是身份竊取，投資者往往無法察覺這些詐騙行為，進而造成資金損失。此外，詐騙行為的增加也會對金融市場的穩定性帶來風險，信任危機可能會蔓延，導致整個金融市場的動盪。

AI 導致了現有的金融監管體系面臨重大挑戰。AI 技術的發展速度遠超過了現有法律和規範的更新速度，傳統的監管方式難以應對日益複雜的 AI 詐騙行為。這需要監管機構加大在 AI 監管和反詐騙技術上的投資，並更新法律框架來加強對 AI 濫用的規範和懲治。

肆、 現有防範措施與監管挑戰

在 AI 技術的快速發展下，衍生出了日益複雜的詐騙手法。詐欺行為結合深偽技術(DeepFake)、自動化社交工程等 AI 應用，提升了假冒的真實性與跨國界性，對全球金融安全構成嚴峻的挑戰。為因應此一趨勢，各國的金融機構以及監管機構已陸續採取多元防範措施，包括多元身分驗證(MFA)、異常交易偵測以及反詐騙模型的建置等技術。然而，在面對快速發展且跨域性的 AI 詐欺，目前的監理制度仍顯不足，會面臨法規無法與時俱進而不符現況、跨國資料共享的困難以及國際合作機制不完善等難題。正因如此，為有效解決全球不斷升級的金融犯罪問題，彌合關鍵的資訊差距，多方合作已成為打擊 AI 詐騙犯罪的關鍵要素之一。國際刑警組織自 2022 年啟動「全球快速支付干預機制」（International Global Rapid Intervention of Payments, I-GRIP）以來，已協助各成員國攔截逾 5 億美元之犯罪所得，顯示跨境合作機制在防堵 AI 詐欺與追回金融詐騙資金上具高度成效與潛力 (林書立, 2024)。本文將進一步探討在金融科技領域中 AI 詐騙之現有防範措施及其監管挑戰。

一、 現有防範機制

(1) 多重身分驗證

在 AI 詐騙迅速演化的情況下，如何建立有效且可信的身分驗證機制已成為金融安全治理的核心。多重身分驗證（Multi-Factor Authentication, MFA）是目前廣泛使用的重要防範機制。核心概念在於結合不同類型的身分驗證因素。身分驗證因素主要可分為三種類型：所知之事(something you know)，又稱知識因素、所持之物(something you have)，又稱持有因素、所具之形(something you are)，又稱生物因素。所知之事如密碼或 PIN 碼，所持之物象是手機裝置、自然人憑證等，而所具之形則是如指紋或人臉辨識、虹膜辨識等。

根據金管會於 2023 年 10 月發布的《金融服務業辦理數位身分驗證指引》[10]明文鼓勵金融業者導入包括生物辨識與行為驗證技術在內的多重驗證機制。該指引指出，在推動金融創新的同時，金融機構也應兼顧用戶資訊與財產安全，並審慎評估數位身分的風險評

[10] 保險相關法規查詢系統 https://law.lia-roc.org.tw/Law/Content?lsid=FL102313

估，確保驗證措施具有實質的防詐效益。此措施指引強調了金融科技應以風險為導向設計身分驗證流程，與法規併行實施，成為防止AI 詐騙的第一道關鍵防線。

目前金融機構已積極導入多重身分驗證技術，特別是在數位金融服務平台上。常見的像是一次性密碼(One-Time Password)、生物辨識技術、以及硬體憑證(如自然人憑證)。生物辨識驗證已被廣泛應用於網路銀行及行動金融應用程式中，利用使用者獨一無二的生理特徵(如指紋、臉部輪廓)做為第二層的驗證，能有效降低帳號遭盜用的風險。

(2) 行為分析與異常偵測

在面對詐騙行為日益多樣化的年代，金融機構除了上述的驗證方式之外，日漸重視行為分析與異常偵測，作為強化防詐的關鍵機制。行為分析指的是依據使用者在數位平台的操作行為，包含滑鼠的點擊軌跡、輸入習慣、以及登入頻率與 IP 位置等，建立了專屬於使用者的行為輪廓。一旦系統偵測到偏離以往的行為模式，例如突然出現海外 IP 的大量交易請求，便自動啟動風險控管程序，觸發額外的身分驗證程序、暫停交易或者是向使用者進行操作確認。此類

偵測具備動態調整能力，能有效的攔截潛在的異常行為。隨著機器學習與數據分析的發展，行為分析模型可透過持續的訓練學習、調整來提升準確度與敏感度，是對抗 AI 詐騙的重要分析工具。

在實務應用上，已經有多家金融機構導入行為分析系統來增強資安防護。以第一銀行為例，於 2024 年 4 月上線了 AI 警示帳戶預警模型[11]。此模型透過自動化分析大量交易數據和用戶行為模式，對網路銀行、ATM、臨櫃交易之行為進行異常機率的推算，針對判定高異常者，進行持有人的審查；而中低異常者，則持續監控並強化身分核實過程。透過系統即時監控及分析攔阻詐騙行為，全面性防堵詐騙案件發生。而另一家匯豐銀行(HSBC)，便採用 BioCatch 所提供的以 AI 為基礎的行為生物識別平台，透過監測用戶與設備互動的細節，例如打字節奏、滑鼠移動軌跡、觸控螢幕操作方式等方式來建立行為輪廓，以進行異常行為辨識 (Arif, 2024)。若是系統偵測到異常操作模式，即便登入憑證正確，仍可觸發額外的驗證流程，以防止帳戶遭冒用。

[11] 第一銀行「AI 警示帳戶預警模型」上線 全方位守護民眾資產安全 https://www.firstbank.com.tw/sites/fcb/touch/zh_TW/1565703032980

隨著科技不斷進步，以 AI 為基底的行為分析與異常偵測系統，不僅僅是輔助工具，而是成為金融資安防護中不可或缺的核心力量。這些系統不但能即時掌握用戶操作行為的細微變化，還能隨著時間自我學習、調整判斷標準，成為防詐的第一道守門員。面對 AI 詐騙手法愈發複雜多變，這類技術提供了更靈敏、更即時的防護機制，使其成為金融機構強化資安架構的重要支柱。

(3) 反詐騙模型的建置

傳統的反詐騙系統多以靜態規則為基礎，例如黑名單比對、交易金額門檻值設定或異常 IP 封鎖等方式。然而，面對詐騙手法日益多樣且複雜，這類靜態規則往往難以即時應對新型態攻擊。因此，金融機構逐步轉向建構動態的反詐騙模型，運用 AI 與機器學習技術來學習正常使用者行為模式，進而偵測潛在的異常行為與高風險交易。此類模型具有自我學習與持續演進的特性，能夠即時調整，使其更貼近真實情境中的變動風險。

根據經濟日報 2024 年 8 月的新聞報導[12]指出，各家銀行為強化阻詐機制，紛紛建立了跨部門小組，並導入神盾[13]、天網 AI[14]、鷹眼模型[15]等自有反詐系統，在 2024 年上半年成功攔阻 38 億元的詐騙資金。以國泰世華為例，不僅成立了反詐騙專案小組，更於 2023 年升級「國泰盾」系統，擴充了犯罪者模組。該模組運用大數據與 AI 技術分析詐騙者常見的操作模式，例如異常約轉設定、登入頻率異常等，進而提取詐騙行為特徵。該系統對可疑帳戶的鎖定精準度已達 90%，異常偵測效率則提升了三倍，顯示 AI 模型在實務上的可行性與成效。

AI 技術在反詐騙模型中的應用持續深化，尤其擅長處理高維度、非結構化資料，如用戶行為、語意訊息、裝置特徵等，成為金融業建構預測型防詐機制的關鍵工具。研究顯示，AI 增強的行為分析能有效檢測釣魚與其他社交工程詐騙中難以察覺的微妙操控線索，顯著降低成功攻擊的發生率（Marsh & McLennan Agency LLC,

[12] 國銀上半年阻詐 38 億元！神盾、天網 AI、鷹眼模型齊上陣 https://money.udn.com/money/story/5613/8154934

[13] 中國信託銀行以 AI 自主監控前／中／後期詐騙行為，結合 KYC、風險標籤與 ATM 通報等機制，有效識別可疑轉帳與車手操作

[14] 元大銀行開發的大數據監控系統

[15] 北富銀與警方共同建立的交易監控模型

2025)。舉例而言，AI 透過自然語言處理（NLP）可自動辨識偽造通知信件的語句異常、排版結構，進一步預測可能的攻擊手法並即時封鎖。這類技術不僅提升整體偵測效率，也大幅降低對人工審核的依賴，有助於縮短反應時間並強化防護能力。

二、 監管挑戰與不足

隨著 AI 在金融領域的應用日益普及，AI 相關的金融詐騙的濫用也成為如今金融業不可忽視的風險。金融以及監理機構雖然積極面對，但是就法律、現實以及技術三個層面，現行的監管制度仍舊存在明顯的不足與挑戰。法律層面面臨了現行法規對新型犯罪的難以界定性的問題；現實層面則涉及跨國犯罪的執法困難、權責界定不明、合作困難等問題；而最後在技術層面上，因 AI 本身的高度複雜度與迅速發展，亦使得傳統的監管手段難以即時因應。以下我將從這三個面向進行探討，分析現階段監管體系在面對 AI 濫用於金融詐騙時會遇到的挑戰與限制。

(1) 法律層面

在法律層面，現行法規針對 AI 技術的規範明顯不足，法規架構呈現明顯的滯後性。多數國家的金融監管體系原本就是為因應傳統金融服務模式所設計，缺乏針對新型的金融服務所可能帶來的風險的評估與具體規範。現行法制面對 AI 的黑箱運作特性以及高度不可預測性，無論在法規的適用與實際執行上，皆存在模糊地帶。 此外，法律的制定與修正曠日廢時，與 AI 科技創新的高速發展形成落差，導致風險管理與責任追究的法律難以即時調整，無法有效回應 AI 應用所帶來的挑戰與問題。

為了因應人工智慧帶來的變革，各國已開始積極制定 AI 法律。歐盟於 2023 年 12 月通過人工智慧法案，為全球首部的 AI 立法。根據風險程度對 AI 技術進行了四類的分級並進行不同程度的監管。若人工智慧技術涉及侵害基本人權或公共利益的高風險應用情境，將被明文禁止使用。其中包括，運用 AI 操控人類行為、影響其自由意志，或者是無特定目的地從網路或監視影像中擷取臉部特徵以建立辨識資料庫，又或者應用含有敏感個人屬性的生物辨識分類系統

（如政治立場、宗教或哲學信仰、種族與性傾向等），此類型的 AI 應用皆在被禁止之列[16]。

目前台灣已積極推動人工智慧相關法規的建構，並著手制定人工智慧相關政策指引，2023 年 8 月行政院公布了行政院及所屬機關(構)使用生成式 AI 參考指引，以規範公部門生成式 AI 的使用原則；同年 10 月，金管會亦發布了金融業運用人工智慧之核心原則，特別針對金融產業導入 AI 提出具體準則[17]。此外，目前行政院及相關部門亦積極研擬制定人工智慧基本法，旨在建立有利於人工智慧技術發展與應用的良善法治環境，並作為引導我國各級機關推動、規範及促進人工智慧應用之原則性依據[18]。

法律的制定與修正曠日廢時，與 AI 科技創新的高速發展形成落差，導致風險管理與責任追究的法律難以即時調整，無法有效回應 AI 應用所帶來的挑戰與問題 。這種法律上的滯後性，也間接加劇了現實層面跨境執法和技術層面模型解釋的困難。

---

[16] 專訪台灣玉山理事陳世杰 淺談 AI 時代之法律議題 https://www.mjtaiwan.org.tw/pages/?Ipg=1007&showPg=1817

[17] AI 時代的十大新型態法律問題（上）https://futurecity.cw.com.tw/article/3341

[18] 人工智慧基本法草案預告 促進創新兼顧人權與風險 https://www.nstc.gov.tw/folksonomy/detail/87e76bcd-a19f-4aa3-9707-ca8927dcb663?l=CH&utm_source=rss

(2) 現實層面

人工智慧應用於金融科技的現實層面，目前仍面臨多重的挑戰，不僅限於技術治理與法規滯後，更深層的牽涉到跨國執法的可行性。AI 詐騙往往跨越國界，舉例來說，犯罪行為可能由單一國家發動，透過多國的金融系統洗錢，最終再回到詐騙集團的所在地。這導致了單一國家的監管與執法機構面臨管不到也抓不著的困境。英國國家打擊犯罪調查局（NCA）局長 Graeme Biggar 也指出，犯罪分子已經開始利用 AI 技術製作勒索軟體和偽造圖像，這使得執法機構面臨新的挑戰。Biggar 強調，這些犯罪行為不再局限於單一國家，而是成為一個全球性的問題，全球須共同面對[19]。

然而，儘管跨國合作對應對這些挑戰至關重要，但在現實中，跨國合作的效果往往受到多重因素的限制。不同國家擁有獨立且相互差異甚大的法規與執法標準，也使得國際間的協作難以達到預期效果。更進一步，資訊落差與數據共享的不一致性，往往使得合作的深度與成效受到限制。此外，一些金融體系通常呈現高度管制與

[19] 壞人也會用 AI！人工智慧將改變犯罪格局
https://www.technice.com.tw/techmanage/infosecurity/77521/?utm_source=chatgpt.com

封閉的結構，對於 AI 技術的應用存有保留，並且內部設計細節涉及商業機密，這無疑進一步加大了監管上的難度。

AI 在金融科技領域的應用雖帶來效率提升，但也暴露出跨國執法困難、權責不明和合作障礙等現實問題。這些現實挑戰，又與法律框架的不完善以及技術的複雜性相互交織，使得打擊 AI 詐騙更顯艱鉅。而為了有效應對這些挑戰，需加強國際合作，完善法律框架，並提升監管機構對 AI 技術的理解和應用能力。

(3) 技術層面

除了法律和現實層面的挑戰，AI 模型本身的技術特性也對監管設計構成重大挑戰 。這些技術上的限制，進一步影響了監管機構在法律框架下的實施和跨國合作的成效。首先，AI 模型模型的黑箱性難以解釋金融決策其決策邏輯，這以增加了金融監管機構難以對其進行審查，因為難以評估模型是否符合法規需求，是否有偏誤並產生了不公平的結果 (Bajracharya, Aakriti & Khakurel, Utsab & Harvey, Barron & Rawat, Danda B., 2022)。此外，AI 模型高度仰賴大量數據進行訓練，若資料本身存在偏誤或代表性不足，亦可能導致錯誤決策。更進一步地，由於 AI 模型具有自主學習的特性，可能在未經監

管核可的情況下，自動改變風險評估準則，進而對金融體系的穩定性構成潛在威脅。因此，面對 AI 技術應用於金融產品與服務的情境，必須建立完善的模型驗證機制，並強化模型的可解釋性，以有效降低潛在風險。

與此同時，AI 技術也被不法分子應用於詐騙行為，例如先前深度偽造技術冒充身分，或者是使用自動化演算法進行大規模詐騙交易。這些新技術讓詐騙行為更加的智能化與具備隱匿性，導致以傳統監管規則為基礎的防詐機制難以偵測，因此，監管機構須持續研發更具即時且與時俱進的 AI 監測工具，以應對快速演化的詐騙技術。

伍、 結論

隨著人工智慧技術在金融產業的快速普及，其在業務效率、風險控管與客戶服務等層面所帶來的正向效益已被廣泛肯定。然而，AI 的高度可擴展性與自動化能力，也同樣被不法分子用於詐騙行為，形成金融科技發展下的新型風險。本報告指出，詐騙者可透過 AI 進行身份偽造、建立虛假交易平台、操控市場訊號，甚至應用深偽技術（Deepfake）偽裝成知名人物或金融機構客服人員。這些高度擬真的詐騙手法使消費者難以分辨真偽，大幅提高了詐騙的成功

率，也使得金融機構在第一線防範上承受前所未有的壓力。AI 在金融詐騙中的濫用已對整體金融市場穩定性與消費者權益構成實質威脅。

雖然各大金融機構已逐步導入如多重身分驗證、行為分析、異常偵測與 AI 反詐模型等技術，並有實際案例顯示成功攔阻詐騙資金，例如國泰世華與第一銀行導入的系統即具有高度成效，然整體監管體系仍面臨三大層面的挑戰。第一，在法律層面，現行法規無法即時回應 AI 帶來的新型犯罪型態，且對黑箱模型的法律適用性存有模糊地帶；第二，在現實層面，詐騙往往涉及跨國操作，使單一國家監管與執法權責難以涵蓋，跨國合作又因法律不一致、資訊共享有其侷限性；第三，在技術層面，AI 模型的不透明性與自我演化能力，使得監管機構難以理解其決策邏輯，也無法及時察覺模型異常所帶來的潛在風險。因此，AI 所帶來的精準詐騙手法，不僅提升了攻擊的隱蔽性，也凸顯現行監管制度在法律、技術與實務層面的侷限性。

因應上述風險，本研究認為應該導入具自我學習的反詐模型與即時異常偵測系統，持續優化模型訓練資料來源與演算法風險控管能力。同時，監管機關應加速 AI 風險法規的制定，建立針對 AI 模

型的審查準則與可解釋性要求，此外，持續推動國際間的合作框架，協助打擊跨境詐騙行為。

未來的研究亦可深入探討 AI 技術的可解釋性（Explainability）、模型治理（Model Governance）與可信任 AI（Trustworthy AI）之建構，以強化技術透明度與法律的可適性。期許未來在享受 AI 科技帶來的效率與創新的同時，能有效控管其潛藏的風險。

陸、 參考文獻

ArifMohammad. (2024 年 8 月 20 日). Biometric Identity Management in Banking: Innovations, AI Integration, and Use Cases. 擷取自 https://www.linkedin.com/pulse/biometric-identity-management-banking-innovations-ai-integration-reinc?utm_source=chatgpt.com

Bajracharya, Aakriti & Khakurel, Utsab & Harvey, Barron & Rawat, Danda B. (2022). Recent Advances in Algorithmic Biases and Fairness in Financial Services: A Survey. 擷取自 ResearchGate: https://www.researchgate.net/publication/364505799_Recent_Advances_in_Algorithmic_Biases_and_Fairness_in_Financial_Services_A_Survey

ButtsDylan. (2025 年 2 月 13 日). Crypto scams likely hit a new record in 2024, driven by ‘pig butchering’ and AI, says Chainalysis. 擷取自 https://www.cnbc.com/2025/02/13/crypto-scams-thrive-in-2024-on-back-of-pig-butchering-and-ai-report.html

Charlotte Thompson&Tiana Kelly. (2023 年 10 月 30 日). Exploring AI Threats: Package Hallucination Attacks. 擷取自 https://www.darktrace.com/blog/when-hallucinations-become-reality-an-exploration-of-ai-package-hallucination-attacks

Marsh & McLennan Agency LLC. (2025). Client Advisory: Social Engineering in the Age of AI. 擷取自 https://www.mcgriff.com/resources/articles/social-engineering-age-of-ai/

SanthanaPrakash. (2025 年 3 月). Industry insight: How generative AI is used for financial fraud detection. 擷取自 https://eandt.theiet.org/2025/03/18/industry-insight-how-generative-ai-used-financial-fraud-detection

Sunny. (2024). 金融業的未來：生成式人工智慧的潛力與應用分析. 擷取自 先行智庫: https://www.kscthinktank.com.tw/blog/金融業的未來：生成式人工智慧的潛力與應用分析/

刑事警察局公共關係室. (2024). 當心虛假名人!AI 深偽詐騙如何騙走你的財富. 擷取自 內政部警政署刑事警察局: https://www.cib.npa.gov.tw/ch/app/news/view?module=news&id=1887&serno=284b7728-56a0-4191-b794-4df960ed3f96

林書立. (2024). 防詐最前線 AI 智慧金融合作分析. 擷取自 https://dsp.twse.com.tw/public/static/downloads/brokerDepartment/%E5%B0%88%E9%A1%8C1-1-%E5%88%91%E4%BA%8B%E8%AD%A6%E5%AF%9F%E5%B1%80%E6%9E%97%E7%A7%91%E9%95%B7-%E9%98%B2%E8%A9%90%E6%9C%80%E5%89%8D%E7%B7%9A%20AI%E6%99%BA%E6%85%A7%E9%87%91%E8%9E%8D%E5%90%88

美商賽仕電腦軟體(SAS). (2019). 金融與風險管理領域的人工智慧應用：認識新世代分析技術並從中獲益. 擷取自 https://www.sas.com/content/dam/SAS/documents/analyst-reports-papers/tw/artificial-intelligence-banking-risk-management-110277-tc.pdf

財團法人金融研訓院. (2024). 113 年資產管理論壇活動摘錄. 擷取自

https://www.sfi.org.tw/files/3868/113%E5%B9%B4%E8%B3%87%E7%94%A2%E7%AE%A1%E7%90%86%E8%AB%96%E5%A3%87%E6%B4%BB%E5%8B%95%E6%91%98%E8%A6%810710.pdf

陳安斌;陳莉貞;蘇秀玲;郭怡君 . (2018 年 1 月). 我國發展機器人理財

顧問之研究. 擷取自

https://webline.sfi.org.tw/download/resh_ftp/AMEDFund/%E6%88%91%E5%9C%8B%E7%99%BC%E5%B1%95%E6%A9%9F%E5%99%A8%E4%BA%BA%E7%90%86%E8%B2%A1%E9%A1%A7%E5%95%8F%E4%B9%8B%E7%A0%94%E7%A9%B6.pdf

謝昀澤. (2022). 當 FinTech 遇上 DeepFake. 擷取自

https://www.fisc.com.tw/Download/102_08.pdf