保險業
前瞻經營講座
紀實115/03/10

國泰金控數數發中心副總經理
劉浩翔副總經理
智能重塑：
從底層邏輯到金融實戰AI治理

一、序言：站在人類歷史的新轉折點

在 2026 年的今天，金融業正經歷一場前所未有的「智能地震」。國泰金控數數發中心副總經理劉浩翔，在講座開場時，首先定義了自己對 AI 的複雜情感 _ 那是「充滿期待又怕受傷害」的混合情緒。擁有近 20 年金融資訊開發與數據應用資歷的他，將生成式 AI 定位為，繼互聯網（Internet）後，人類歷史上最具顛覆性的科技突破點。

劉副總強調，這不只是單純的技術升級，而是一場跨越部門藩籬、重塑企業戰略的全面變革。從產險、壽險到銀行，從後勤稽核到前線業務，AI 已成為一種共同語言。他帶領同學從 AI 的「機率本質」出發，穿梭全球金融巨頭的實戰現場，最後歸於最關鍵的「治理邊界」。本紀實將從底層邏輯出發，解構 AI 代理人的實戰應用，並深入探討金融業最核心的治理與風險控管課題。

二、第一樂章：解密引擎_從預測到創造的本質躍遷

(一)典範轉移：數據收斂 vs. 文字接龍

劉副總精確定義了生成式 AI 與傳統機器學習（ML）的本質差異，關於傳統機器學習，是基於歷史數據進行「預測與分類」，追求的是數據的收斂；現在的生成式 AI，本質是「理解與創造」，其運作核心是「下一個 token 的機率預測」。

他強調，大型語言模型（LLM）並不具備真實意識，而是一台擁有兆級參數、高維度數學處理能力的「超級文字接龍機器」。AI 不在意事實真相，它在意的是根據上下文，下一個字出現的最高機率是多少。

(二)Transformer與自注意力機制

AI 展現推理能力的關鍵在於 Transformer 架構。包含「全景視野」不同於傳統逐字閱讀（RNN），Transformer 具備一次掃視整頁內容的能力；以及「自注意力機制 (Self-Attention)」，這讓 AI 在處理長文本時，能精準標記「誰與誰最相關」，捕捉複雜的上下文脈絡與因果關係。

(三)Scaling Law 與模型生態

劉副總提到，目前模型性能遵循 Scaling Law，透過堆疊算力與資料量提升智力。然而，隨之而來的是邊際效益遞減，競爭焦點已從純粹的「智力」轉向「穩定性」與「企業對接」。企業在選擇模型時，常在追求極致效能的「閉源巨頭」（如 OpenAI、Anthropic、Gemini）與具備佈署彈性的「開源勢力」（如 Llama、DeepSeek）之間取得平衡。

三、第二樂章：自主行動 _ AI Agent (數位員工) 實戰架構

劉副總強調，AI 應用的下一個主戰場是從「說」走向「做」的 AI Agent (代理人)。這不再只是對話，而是具備目標導向與持續行動能力的數位員工。

保險業
前瞻經營講座
紀實115/03/10

國泰金控數數發中心副總經理
劉浩翔副總經理
智能重塑：
從底層邏輯到金融實戰AI治理

他提出AI Agent 的三層核心架構，第一層是感知層 (Perception)，賦予 AI 雙眼與雙耳。它能讀取 PDF 財報、辨識理賠影像，甚至分析客戶情緒；第二層是大腦層 (Brain)，由 LLM 結合思維鏈（Chain-of-Thought）組成。AI 能將複雜任務拆解為執行步驟，並展現推導過程；第三層是執行層 (Action)，指透過 API 或MCP調用工具，相當於AI的手功能，讓AI 能主動登入系統、比對名單，並完成發送 Email 等執行動作。他提出一個重要的的現實，在個人應用場景，AI Agent有80% 的任務達成成功率或許夠用，但對於企業而言，需要的是穩定在 90% 或以上的任務完成度才算可靠。這需要搭配審計、迴圈（Rollback）與重試機制，並落實「人機共治」，這任務成功率，是AI Agent技術未來能否邁入大規模企業級應用的關鍵點。

四、第三樂章：全球戰場 _ 金融與保險業的轉型案例

GenAI 正在重塑金融價值鏈的每一個節點，從行銷、核保到客服。他以效率革命與保險實務分別說明。

(一)效率革命的標竿案例

1.Klarna (支付)：其 AI 代理人處理了 2/3 的客服對話量，工作產能相當於 700 名全職客服（FTEs），問題解決時間從 11 分鐘大幅縮短至 2 分鐘。

2.摩根士丹利 (Wealth Management)：打造 AI 助手，秒級消化十萬份研究報告，賦予理專精準決策能力。

3.摩根大通 (JPMorgan)：推出 LLM Suite 知識助手，提升分析師研究報告的生成效率。

(二)保險實務深耕案例

1.蘇黎世保險 (Zurich)：提取非結構化醫療報告數據，縮短核保壓力。

2.Allianz：建構內部全球知識庫，回答公司流程與保單問題，打破資訊壁壘。

3.Lemonade (數位保險)：從投保到理賠完全交由 AI 處理，實現極致的「秒級體驗」。

4.東京海上：利用 AI 進行複雜的「巨災風險預測」。

五、第四樂章：駕馭巨獸 _ AI 治理四大關鍵風險深潛

在金融業運用生成式 AI 的實踐中，針對 AI 風險的評估與治理通常聚焦於四大核心領域，以下整理這四大風險情境與對應的治理考量：

(一)資料與隱私風險 (Data and Privacy Risk)

考量員工可能不慎將包含客戶個人識別資訊（PII）的理賠卷宗或敏感文件上傳至未受保護的公開外部大型語言模型（LLM）平台進行摘要或處理，進而導致嚴重的資料外洩。

企業要評估是否建立地端模型（On-premise）或專屬的企業級雲端 API，以確保數據不外洩。在資料餵入 AI 模型前，應做好適當的加密機制。並實施嚴格的基於角色的存取控制（RBAC），確保 AI 的權限不超過操作該 AI 的員工。

保險業
前瞻經營講座
紀實115/03/10

國泰金控數數發中心副總經理
劉浩翔副總經理
智能重塑：
從底層邏輯到金融實戰AI治理

(二)模型風險與幻覺 (Model Risk and Hallucinations)

因為AI 在生成理財建議或解釋保單條款時，可能憑空捏造不存在的條文（幻覺），導致不當銷售（Mis-selling）或引發客訴。

所以，在治理考量因素上，我們不能單純依賴 LLM 的原生記憶與知識，可採用檢索增強生成（RAG）架構，將 AI 的回答，錨定在企業內部的標準答案上；同時，建立「人機協作」（Human-in-the-loop）機制，所有涉及決策（如拒保、拒賠）的 AI 輸出必須經過人類專家覆核。

(三)系統性與資安風險 (Systemic and Information Security Risk)

為避免惡意使用者可能透過特製的提示詞攻擊（Prompt Injection），繞過 AI 系統的保護，誘使 AI 洩漏敏感指令、吐露後台資料或做出不當承諾。

我們得針對第三方模型服務（如服務停機）建立營運持續計畫（BCP），並建置「輸入端與輸出端護欄」（Input/Output Guardrails），主動攔截惡意指令與違規輸出。

(四)公平性與演算法偏見 (Fairness and Algorithmic Bias)

因為，AI 模型若過度依賴具偏見的歷史數據，可能對特定年齡、性別或職業族群產生系統性的歧視（如拒保偏見）。

就得進行訓練資料的多元性與代表性檢測，避免使用人類生理特徵作為模型訓練的變數，並執行模型公平性檢測。

六、結論：從數位轉型走向 AI 轉型

演講尾聲，劉副總提出導入AI不能只靠想像，企業成功導入 AI 有「黃金三角」，並提醒企業「知道AI的極限，才能真正駕馭它」。首先，要有專屬數據資產：演算法可以複製，但企業累積的專有數據才是無法取代的護城河，也決定AI知識的廣度；接著，落實人機協作(Human-in-the-loop)，保持人類在決策迴圈中，保有最終判斷以確保專業與合規；最後，要保持迭代實驗文化，不求一步到位，而是從小地方「概念驗證」（Proof of Concept, PoC)做起，評估其可行性與技術價值的過程，具體累積成功經驗後再行擴張。

他指出，未來職場的關鍵技能是 提示詞設計 (Prompting)。這是一場從底層邏輯、業務流程到人類技能的三層轉型。

這場講座不僅是一次技術的解密，更是一份金融從業者在智能時代的生存指南。劉副總強調，引領變革的起點不是演算法，而是建立「信任」。當企業具備了強壯的安全帶，才能在 AI 的高速賽道上，無所畏懼地全速奔馳。

本紀實由淡江大學風險管理與保險學系碩士在職專班
系友楊孝翔撰文
國泰金控數數發中心劉浩翔副總經理同意刊登

劉浩翔副總經理(前排左二)、何佳玲主任(前排左三)與同學合影