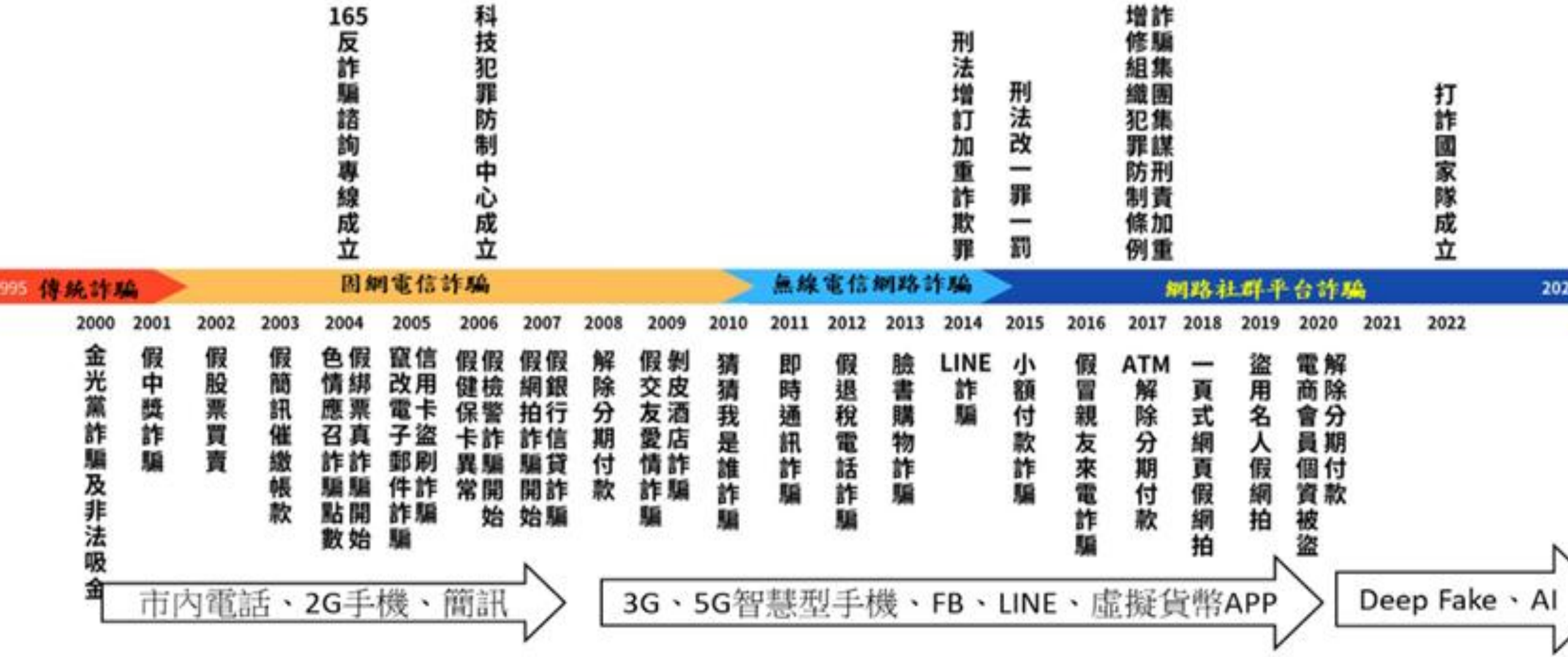
AI科技詐騙犯罪預防與因應

主講人:范振家　　博士

電信網路詐騙演進時序表



資料來源:刑事警察局　研究者整理

CONCEPTS

概念界定：先釐清「工具、內容、行為、犯罪」四個層次

在反詐討論中，AI、生成式AI、深偽、詐騙犯罪與 AI 科技詐騙常被混用。若名詞界定不清，政策工具就容易失焦：有些問題屬於技術治理，有些屬於刑事法與金融監理。

- **AI**：能從輸入資料推論並產生輸出的機器系統。
- **生成式AI**：AI 的一支，重點在生成文字、聲音、影像與影片。
- **深偽**：利用AI生成或偽造高擬真人臉、聲音、影像。
- **詐騙犯罪**：以詐術使人陷於錯誤並交付財物或利益的犯罪。
- **AI科技詐騙**：以AI放大詐騙的規模、速度、擬真度與自動化程度。

AI（人工智慧）

重點在「推論」與「輸出」：系統可根據資料與目標，產生預測、建議、分類或內容。

生成式AI

不是所有 AI 都會生成內容；生成式AI特別擅長生成文字、圖像、聲音、影片與程式碼。

深偽（Deepfake）

屬於高風險應用型態，常用於冒充主管、親友、檢警或名人，提高受害人信以為真的機率。

詐騙犯罪

核心不是「有沒有用新技術」，而是是否以不實資訊、冒名、隱匿真相等方式，使人做出錯誤財產處分。

政策判讀原則：AI 是工具，詐欺是行為；AI 科技詐騙則是「以 AI 強化詐欺效果」的複合型風險。

資料來源：OECD AI Principles(2024)；刑法第339、339-4條；詐欺犯罪危害防制條例

AI BASICS

何謂 AI？從人工智慧到生成式 AI 的基本理解

依 OECD 近年更新後的定義，AI 系統是：在明示或默示的人類目標下，能從輸入推論如何產生輸出的機器系統，並可生成預測、內容、建議或決策。

機器系統 不等於一般軟體自動化；其核心在資料運算與推論。

人類目標 AI 仍由人設定目的、限制與使用情境。

推論能力 AI 會依輸入資料推估輸出，而非單純照表操課。

輸出結果 可表現為預測、推薦、分類、生成內容或輔助決策。

AI ≠ 聊天機器人

推薦系統、風控模型、臉部辨識、語音辨識與自動翻譯，都可屬 AI 應用。

生成式AI是子集合

它特別擅長生成語句、圖片、語音與影片，因此更容易被用來製造看似真實的詐騙素材。

深偽的政策意義

深偽不是另一種獨立犯罪類型，而是一種高擬真偽造方式。當它被拿來冒充親友、主管、檢警、名人或官方通知時，就會提高詐騙訊息的可信度與即時壓迫感。

治理含意：不能把所有數位風險都籠統稱為 AI；應區分「推論型 AI」與「生成型 AI」的治理要求。

資料來源：OECD AI Principles(2024) 與 2024 explanatory memorandum

AI TYPES

「推論型 AI」與「生成型 AI」的差異

兩者都屬於人工智慧，但功能焦點不同：推論型 AI 著重分析、判斷與預測；生成型 AI 著重產生文字、圖像、語音、影片等新內容。在反詐治理上，前者多用於偵測與預警，後者則更常被用來提高詐騙素材的擬真度與擴散效率。

核心任務 推論型做分類、預測與評估；生成型做文字、圖像、語音與影片生成。

典型輸出 前者輸出分數、標籤、排序；後者輸出原本不存在的新內容。

常見應用 前者見於風控、異常交易偵測；後者見於聊天機器人、深偽與語音合成。

詐騙風險 前者強化精準鎖定；後者強化冒名、假訊息與假影音擬真。

推論型 AI

依資料特徵與模型結果進行分類、預測、評估或排序。常見於金融風控、信用評分、帳戶異常警示與犯罪預警。

生成型 AI

依提示自動生成文字、圖片、語音、影片或程式碼。常見於聊天機器人、深偽、語音克隆與自動內容生產。

在反詐治理中的意義

推論型 AI 有助於銀行、平台與執法機關及早辨識高風險帳戶、訊息與交易；生成型 AI 則會降低詐騙內容製作成本，提升假冒公務員、親友、客服與投資導師的擬真程度。治理上應分別處理「誤判與透明性」以及「內容真偽與來源標示」兩類問題。

快速判準：輸出若是分類結果、風險分數或預測，多半屬推論型；若是文字、圖像、語音、影片等新內容，多半屬生成型。

FRAUD TERMS

何謂詐騙犯罪？AI 科技詐騙並未改變本質，但大幅放大風險

從刑事法角度看，詐欺的核心在於：以詐術使人陷於錯誤，進而交付財物或取得財產上利益。AI 的加入，改變的是手法表現，不是犯罪本質。

1 施用詐術 例如冒名公務員、假投資平台、假客服、假中獎通知。

2 使人陷於錯誤 被害人相信其為真，降低查證與防備。

3 作成財產處分 匯款、交付帳戶、提供 OTP、交出提款卡或個資。

4 發生財產損害 金錢被轉走，或形成後續擴大損失。

5 具有因果關聯 詐術、錯誤、交付與損失之間具連續關係。

傳統詐騙

主要仰賴話術、人海撥打、紙本文件或低階假網站。

AI 科技詐騙

加入自動生成、聲音複製、換臉、分眾投放與腳本自動化。

結論：AI 不是新的犯罪本質，而是新的犯罪放大器。它讓冒名更像真、觸達更大量、操作更自動、金流更快速。

若冒用政府機關、透過電子通訊或網路對公眾散布，甚至以電腦合成不實影像、聲音犯案，可能涉及加重詐欺等更重法律效果。

資料來源：刑法第339、339-4條；臺北地檢署法律宣導；法務部打詐四法說明

治理導向簡報

AI科技詐騙犯罪預防與因應

以台灣治理實務、法制回應與跨部門聯防為核心

政策重點不再只是「識詐」，而是同步推進平台治理、金融阻詐、深偽查證、跨境執法與被害保護。

治理觀測

61.1萬

截至2026/03/31月均疑似詐騙訊息通報量

24.9萬

經審查確認為詐騙並通報平台下架

530件 / 日

2026年3月平均受理詐騙案件

另同期平均財損約 2.76 億元 / 日；網路購物與假檢警視訊均為高風險場景。

從技術防詐走向制度防詐

重點涵蓋：風險態勢、犯罪鏈、法制回應、應變SOP與政策建議

風險態勢：AI讓詐騙由零星欺騙轉為規模化產業

關鍵訊號

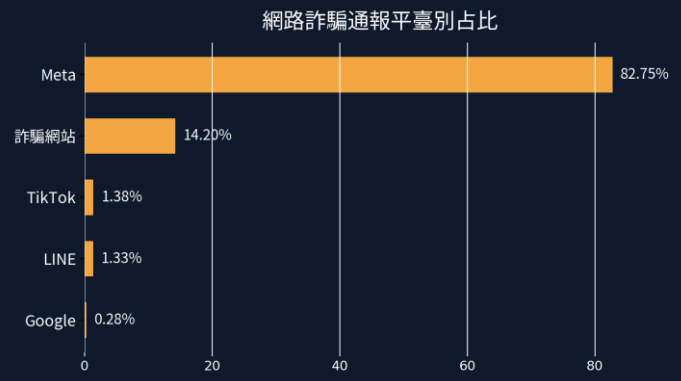
7%
全球僅7%反詐專業人員認為組織已相對充分準備好面對AI驅動詐欺

20%
Entrust指出，生物特徵詐欺中約五分之一涉及深偽

82.75%
台灣最新週報中，Meta仍占確認詐騙訊息最大宗來源

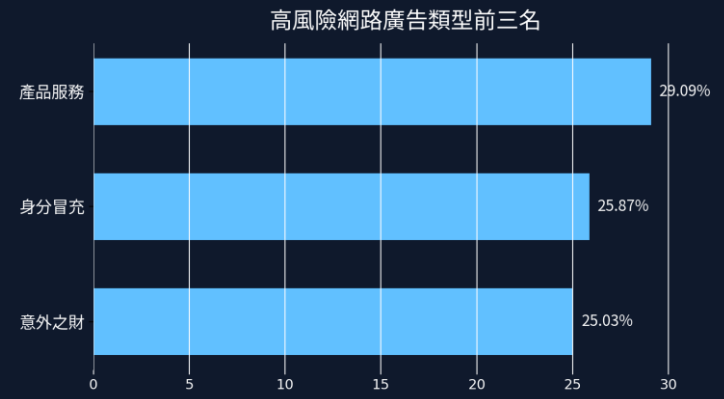
意涵：AI降低詐騙製作成本，卻抬高驗證成本；風險正由個人判斷失誤，轉為整體治理能力不足。

平台別占比



社群廣告與假網站仍是主要入口，表示治理重點必須前移至「投放、導流、驗證」三個節點。

廣告類型前三名



- 產品服務類型高居首位，反映「限時、免運、低價」仍具高誘引力。
- 身分冒充位居第二，AI合成影像與聲音使名人、公務機關、親友冒充更具可信度。
- 投資與意外之財類型則與高報酬、快速回本等心理誘因緊密連動。

MECHANISM

AI改變詐騙經濟學：低成本、高擬真、跨通路、自動化

01 生成腳本與在地化話術

大型語言模型可快速產生多語版對話腳本、客服回覆與「權威」說詞，降低詐團訓練門檻。

02 深偽影音提升可信度

聲音複製、AI換臉與偽造視訊背景，使假檢警、假主管、假親友的即時互動更具說服力。

03 多通路導流與精準分眾

從社群廣告、簡訊、通訊軟體到假網站，攻擊可跨平台串接，並依受害者輪廓動態調整。

04 即時支付與虛擬資產洗白

快速轉帳、第三方支付與USDT等虛擬資產讓金流斷點縮短，追索難度上升。

05 詐騙服務化（Scam-as-a-Service）

工具、腳本、投放、洗錢與客服可分工外包，形成跨國化、模組化的犯罪供應鏈。

分析判斷

AI並未改變詐騙的本質，改變的是「規模、速度與逼真度」。因此，單靠個人警覺已不足，必須改以制度化驗證與跨機關聯防補位。

治理含意：一旦舊式判斷指標（例如錯字多、畫質差、來電口音怪異）逐步失效，政策重點就必須從「內容辨識」轉向「身分驗證」與「交易摩擦設計」。

ECONOMIC VIEW

從犯罪經濟學看台灣反詐「越打越多」

將詐欺視為一個具需求、供給與中介設施的地下市場

詐欺不是單純治安問題，而是可模組化供給、可跨境洗錢、可持續擴張的犯罪市場。若治理只壓制單一手法，而沒有改變利潤結構、需求脆弱性與平台 / 金流基礎設施，供給端就會在有利可圖的條件下持續補位與重組。

治理重點應從「抓多少人」擴展為「如何壓縮詐欺市場的再生能力」。

核心觀點

詐欺市場化後，僅靠個案打擊不足，必須同步處理需求端、供給端與洗錢 / 平台中介。

為何「越打越多」

多數治理壓制既有手法，卻未同步摧毀獲利結構、需求基礎與中介基礎設施。

模型目標

建立兼顧需求端、供給端、中介設施與執法成本的詐欺誘因分析模型。

政策方向

區分「戰術打擊」與「戰略治理」，並納入虛擬資產、國際金融與教育信任修復。

整合自附件《反詐越打越多：犯罪經濟學分析框架》

MODEL

詐欺誘因分析模型：需求端 × 供給端 × 中介基礎設施

用犯罪經濟學解釋詐欺市場為何持續擴張

需求端 | 被害人脆弱性

投資、借貸、求職、交友、消費等需求持續存在。

焦慮、孤獨、貪利與資訊不對稱，會提高可操弄性。

只要社會中存在穩定可利用需求，市場就不缺目標。

供給端 | 詐團獲利性

高報酬、低邊際成本、失敗可快速重來。

話術、名單、平台、AI內容都可複製擴張。

模組化分工使組織即使被打擊，仍可迅速重組。

中介基礎設施 | 平台與金流

社群平台、廣告系統、通訊軟體、假App，提供導流與接觸面。

人頭帳戶、地下匯兌、虛擬資產、場外交易，讓詐款可分層與跨境移轉。

詐欺擴張力 ≈ 需求端脆弱性 × 供給端獲利性 × 中介可得性 ÷ 執法嚇阻強度

供需模型有助於把詐欺視為市場，而非單一案件集合。

WHY MORE

為何會出現「越打越多」？

戰術性壓制不等於結構性減量

01 打掉舊手法，未必打掉市場

電話詐騙受壓制後，詐團可迅速轉向社群平台、假投資 App、交友養套殺與 AI 冒名。

02 供給端已產業化

東南亞詐騙網絡已形成話術、洗錢、擔保服務、虛擬資產與人口販運等完整服務鏈。

03 需求端具有再生性

投資焦慮、情感需求、求職壓力與數位金融新手，持續構成可被切分與操弄的目標市場。

04 執法與犯罪成本不對稱

一次成功詐騙的收益可觀，但政府跨境偵查、凍結、返還與修復成本極高。

「越打越多」的本質，是犯罪供給能力的升級速度，高於治理體系改變市場誘因的速度。

若不處理利潤結構與需求基礎，戰術成功仍可能被市場替代效應抵消。

GOVERNANCE

戰術打擊 vs. 戰略治理

未來政策設計需從個案處理升級為市場治理

戰術打擊｜處理當前案件

- 查緝機房、車手、水房
- 凍結人頭帳戶與可疑金流
- 下架假廣告、封鎖假帳號
- 即時攔阻高風險匯款

功能：提高犯罪者直接成本，限制當前案件的發生與擴散。

戰略治理｜改變市場誘因

- 降低需求端易感性：識詐列課綱、金融素養、平台識讀
- 削弱中介基礎設施：平台責任、VASP 監理、場外交易治理
- 提高跨境洗錢成本：司法合作、資產追徵、數位證據交換
- 修復社會信任：可信識別機制、制度透明與教育

功能：讓供給端不再那麼值得進場。

戰術處理事件；戰略改變生態。

FINANCE × TRUST

虛擬貨幣、國際金融與教育信任修復

超越傳統治安思維的跨領域政策框架

虛擬貨幣不是單一工具，而是基礎設施

它連接詐欺、跨境資金移轉、場外交易、殼公司與地下金融，使詐騙更易跨境擴張與資金分層。

國際金融合作是必要條件

若缺乏 AML/CFT、VASP 監理、可疑錢包風險分級與跨境證據交換，供給端仍能透過全球金流鏈再生。

教育的任務

不只是提醒「不要被騙」，而是訓練民眾：

- 見高報酬先查證
- 接到權威通知先回撥
- 看到感情與金流混合先暫停
- 對平台介面與假 App 保持懷疑

信任修復的任務

未來反詐應納入：

- 平台與公部門可信識別
- 可驗證的客服與通知機制
- 對被害人非污名化支持
- 讓制度重新成為可信訊號

結論：反詐若提升為金融、平台、教育與信任治理，才可能壓低整體市場再生能力。

SCENARIOS

我國高風險場景：AI通常不是單獨作案，而是嵌入既有詐騙劇本

01	網路購物 / 一頁式廣告	以低價、免運、限時促銷導流至仿冒頁面，再由假客服進行二次詐財。
02	假投資 / 名人成分背書	結合AI生成素材、假績效截圖與社群投放，擴大陌生投資的信任假象。
03	假檢警 / 視訊筆錄	偽造警局背景、制服、公文與視訊紀錄，利用權威壓力迫使受害人隔絕外部求證。
04	假冒親友 / 假主管語音	透過聲音複製或即時語音合成，要求匯款、代墊費用或變更付款流程。
05	戀愛交友 / 虛擬資產	使用深偽頭像、AI聊天與假交易平台，讓情感關係與投資誘因同時發生。

近期觀察

- 警政署指出，2025年7月網路購物詐騙平均每日受理90件，較前月增加21%。
- 同年1至7月假檢警詐騙共受理3,114件、財損62.05億元，且近期出現假視訊筆錄。
- 官方周報亦顯示熱門關鍵字常圍繞「特價、免運」，反映消費場景仍是最大入口。

判讀原則

凡同時出現「權威壓力 + 時間急迫 + 封閉求證 + 即時付款」四要素者，即應視為高度風險。

CASE STUDY

案例補充：台北護理師遭假檢警騙走379萬元

1 假郵局來電

謊稱證件遭冒用、普發現金可能被冒領。

2 假警員加LINE

以「視訊筆錄」要求出示身分與帳戶資料。

3 假檢察官恐嚇

宣稱涉詐、遭開拘票，並要求偵查不公開。

4 房貸增貸與交卡

增貸226萬元，交出存簿、提款卡供所謂監管。

5 再索保證金

再騙138萬元，帳戶合計379萬元遭提領一空。

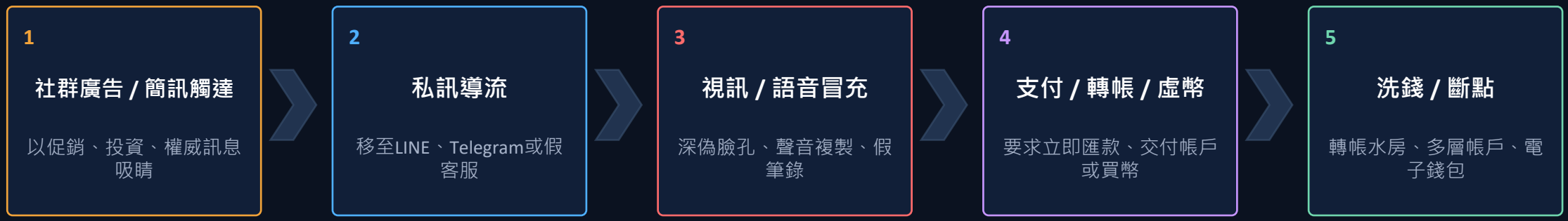
案例顯示的五個警訊

借政策議題製造急迫感	主動轉介公部門形成可信鏈	通訊軟體辦案與視訊筆錄	以偵查不公開隔絕求證	要求交付帳戶工具與金錢

政策意涵：此案顯示，AI不必然單獨構成詐騙核心，但會強化話術生成、視訊擬真與權威冒充效果；真正有效的防線，在於第二通路查證、金融摩擦與公權力識別教育同步前移。

ATTACK CHAIN

典型犯罪鏈：從廣告投放到跨境洗錢，風險節點可以被制度化切斷



對應防線

廣告實名與高風險素材查核	跨平台可疑連結即時標記	第二通路回撥查證機制	高風險轉帳延遲與人工關懷	跨境執法、凍結與追贓

設計原則：治理上不必期待「完全辨識所有假內容」，而應在每一個交易節點增加最低必要之驗證與摩擦。

GOVERNANCE RESPONSE

我國法制與治理回應：從打詐四法走向2.0版綱領與跨業聯防

法制基礎

- 《詐欺犯罪危害防制條例》已於2024年7月31日公布施行，形成專法基礎。
- 「打詐四法」整合詐欺、洗錢、刑訴與通保等面向，回應偵查、下架、扣押與量刑問題。
- 高檢署說明指出，運用電腦合成或其他科技方法製作不實影像、聲音而詐欺者，屬加重處罰重點。

行政院綱領2.0

新增「防詐」面向，並將政策重點前移至：

- 強化AI防制
- 深化跨境合作
- 監管防詐產業
- 強化被害人保護

同時要求：電信實名、平台資訊揭露與管理、金融機構管理警示帳戶與虛擬資產AML、地方政府共同執行。

政策判讀

現行框架已從「宣導為主」走向「治理工具 + 市場義務 + 跨境合作」的複合型模式。

跨業聯防實務

- 金管會於2024年推動「科技防詐」主題式活動，結合PET、合成資料與AI進行聯合學習。
- 聯合自主實證涵蓋銀行、電信、科技公司與國家資安研究院，強化支付、匯款與理賠詐欺預警。
- 金融資安韌性發展藍圖則把應變與復原能力納入下一階段監理重點。

POLICY LEVERS

政策設計重點：政府端宜從五個治理槓桿同步下手

1

身分與內容真實性

建立高風險AI內容之標示、查核與申訴機制；要求廣告主、代操與平台負起最低驗證義務。

2

金流與交易摩擦

對高風險轉帳、首次收款對象、異常跨境支付與虛擬資產轉換導入延遲、提醒與人工覆核。

3

資料共享與隱私保護

以PET、聯邦學習、合成資料等方式共享可疑樣態，在不過度揭露個資下提升預警準確度。

4

地方治理與分眾識詐

針對長者、青少年、小微商家與公部門採購承辦等高風險族群設計差異化教材與演練。

5

跨境偵查與追贓

將電子證據保全、跨境資料交換、虛擬資產追蹤與被害通報整合為常態化合作流程。

治理邏輯：不是把責任全部推回給使用者，而是將平台、支付、電信、地方政府與執法機關納入同一風險治理鏈。

RESPONSE SOP

機關與組織應變SOP：被騙當下的處置速度，決定可否阻斷財損

停 停止匯款 / 停止對話

立刻中斷任何付款、面交、帳戶交付或下載遠端控制APP。

查 改用第二通路查證

自行撥打官方電話、銀行客服或165，不使用來訊附帶的連結與號碼。

通 內部通報與對外求助

公務機關與企業應立即通報資訊、法遵、主計或主管鏈，不可單線處理。

凍 圈存 / 凍結可疑帳戶

盡速向165、銀行或警方提供交易資訊，爭取暫時凍結與阻卻轉出。

存 完整保全數位證據

保存對話紀錄、匯款證明、網址、假視訊截圖、來電號碼與裝置紀錄。

說 啟動受害溝通與復原

避免羞愧沉默造成二次被害；同步通知家屬、財務窗口與可能受影響單位。

實務提醒：切勿另行委託網路上聲稱可「代追回款項」之陌生人，以免遭遇二次詐騙。

結語

面對AI詐騙，
治理的核心不在於追逐每一種新工具，
而在於建立可驗證、可攔阻、可復原的制度。

三個結論

- 第一，AI使詐騙的「規模、速度、擬真度」同步升高，傳統識別經驗正快速失效。
- 第二，台灣已形成法制與跨業合作基礎，下一步關鍵在於落地執行與量化治理。
- 第三，真正有效的防詐，必須把個人警覺、組織流程、平台責任與國家能力串成一體。