星球永續健康線上直播

智慧數位資安 (2)

AI 時代數位資安策略

2026 年 4 月 8 日

在當前數位化與智慧化時代資安已不再僅是防禦機制的優化，而是隨著全球數位化與新秩序快速重塑與持續變動呈現前所未有的動態與複雜特徵。數位資安核心亦由單點防護，轉向建立完整的生命週期管理架構，涵蓋風險辨識(Identify)、系統保護(Protect)、威脅偵測(Detect)、事件回應(Respond)與系統復原(Recover)，強化數位資安韌性，使系統在遭受攻擊後仍能迅速恢復並維持穩定運作。本週我們將探討智慧數位資安生命週期的關鍵架構，並以金融領域中的資安實務案例說明。

健康科學新知

中東衝突僵持 國際勢力關注:「戰略模糊」

當前中東衝突同時在戰場、外交、能源體系與國際輿論中的全面性對抗。伊朗總統佩什基安向美國與全球發出公開信，強調伊朗不是安全威脅，也不是主動發動戰爭的一方，而是一個歷史悠久、以自衛為主的文明國家，試圖塑造伊朗形象。美國則在川普主導下對伊朗展現出強硬且矛盾的態度。川普一方面揚言奪取伊朗石油、控制哈格島，威脅摧毀伊朗的發電設施與油井，顯示美方已把能源設施與海上通道視為戰略核心。另一方面持續釋出美伊仍可透過第三方接觸與談判的訊號，使美國的對外政策呈現出一種軍事威脅與外交施壓並行的戰略模糊。美軍正為可能的地面行動做準備，伊朗方面則強烈警告美軍地面軍事行動將遭遇直接而猛烈的反擊，顯示雙方都已為更大規模的對抗預留空間。這場衝突的影響已外溢中東地區。葉門胡塞武裝加入戰局，黎巴嫩南部戰線升高，波灣地區的電力設施與海水淡化廠也開始成為攻擊目標，顯示戰爭正從軍事設施、港口與油井，進一步延伸到民生基礎建設。由於荷莫茲海峽與紅海航道都是全球能源運輸的重要節點，任何封鎖或打擊都會迅速影響國際油價與供應鏈穩定，因此這場衝突已不只是地區安全問題，也成為全球能源與經濟風險來源。對美國-以色列發動之軍事行動上

週 100 多位國際法專家簽署了公開信，對美國、以色列和伊朗在中東戰爭中嚴重違反國際法的行為表示深切關注。

全球航運關鍵節點:「咽喉命脈」

全球海上貿易對經濟體系仍具關鍵地位，即使在多元運輸方式發展下，航運仍承載約 85%的貨物流量。然而，當前多重風險正在削弱其穩定性，包括地緣政治衝突、非國家武裝力量的科技能力提升，以及氣候變遷對關鍵航道的影響。特別是荷姆茲海峽的封鎖，使全球約五分之一的石油與液化天然氣出口受阻，顯示海上咽喉點對能源供應的高度敏感性。全球多個戰略性海峽與運河構成現代貿易網絡的核心節點，例如馬六甲海峽、蘇伊士運河與巴拿馬運河等。一旦這些節點受阻，即使存在替代路線，也往往需要大幅繞行，導致運輸時間與成本顯著增加。不同航道的重要性並不僅取決於流量，亦取決於是否存在替代通道，例如荷姆茲與土耳其海峽因缺乏替代路徑而更具脆弱性。無人機與飛彈等低成本技術已使武裝團體能威脅遠距海上航運，而區域戰爭也容易外溢至海域，干擾糧食與能源運輸。未來若爆發大國衝突，例如美中因台灣議題產生軍事對抗，可能出現封鎖與對商船攻擊，對全球供應鏈造成更劇烈衝擊。同時，各國開始重新評估對海上通道的依賴，並透過軍事部署、能源管線建設與基礎設施投資來降低風險。氣候變遷亦改變航運格局，例如巴拿馬運河因乾旱限制通行量，而北極航道因冰層消退逐漸浮現，使新的戰略節點受到關注。在此背景下，國際物流體系展現一定調適能力，例如改道非洲或發展陸運替代方案，但其容量有限，難以完全取代海運功能，且成本上升已反映在運價與能源價格之中。

伊朗戰爭重塑能源版圖:「能源震盪」

能源危機源於中東衝突升溫，導致荷姆茲海峽關閉，使全球約五分之一、每日約 2,000 萬桶的油氣運輸受阻，對全球能源供應造成即時且顯著的衝擊。短時間內，油價於亞洲、歐洲與美國同步上升，航運保險成本大幅攀升，部分保險公司甚至暫停承保；同時，多家能源企業啟動不可抗力條款，進一步推升市場不確定性，顯示能源市場對地緣政治風險具有高度敏感性。在需求端，能源高度依賴進口的亞洲國家首當其衝。由於

約 80%經該航道輸出的能源最終流向亞洲，日本等國已啟動戰略石油儲備並採取節能措施以緩解短期壓力。然而，此類應對手段具有明顯時效限制，難以支撐長期供應中斷風險。此外能源供應鏈的中斷亦進一步影響肥料、化工與製造等關鍵產業，擴大經濟衝擊範圍，突顯全球能源體系對單一航道的高度依賴與結構性脆弱。面對此衝擊，各國需從結構層面進行調整，包括推動能源運輸路徑多元化、投資替代基礎設施、強化供應鏈韌性與跨國協調機制，以降低關鍵瓶頸所帶來的系統性風險。同時，加速再生能源與低碳能源發展，已成為降低地緣政治依賴的關鍵策略。

日本外交軍武雙重措施:「樞紐崛起」

2026 年春季的東亞安全局勢呈現日本戰略地位快速上升與區域秩序重組同步推進趨勢。法國總統馬克宏訪問日本與南韓，顯示歐洲主要國家正積極強化與日本的安全合作，並在印太布局中展開彼此競逐。對法國而言，日本已成為連結全球安全、供應鏈韌性與印太權力平衡的關鍵節點。馬克宏此行不僅深化法日關係，也回應德國近年積極經營日本的作為，反映出日本已成為歐洲國家爭取亞洲影響力的重要舞台。同時法國也試圖在印太與歐洲兩條戰略軸線上同時維持存在，展現其跨區域大國的角色。另一方面，日本自身也正在推動深刻的安全政策轉型。2026 年 3 月，日本正式部署首批長程飛彈，並同步發展高超音速滑翔武器與戰斧巡弋飛彈部署計畫，顯示其軍事能力正由傳統近距離防禦，轉向具備遠距打擊與反擊能力的新架構。然而，此一軍事轉向也引發地方居民、學者與批評者的疑慮，因為飛彈部署不僅提高地方社會的安全風險，也觸及和平憲法、國家定位與政治正當性的爭議。周邊國家對此亦高度敏感傾向將日本軍事能力提升與歷史記憶、政治右傾化及區域權力競爭連結解讀。

AI 算力擴建跨域合作並起:「算力競權」

人工智慧已從單一技術工具成為重組全球產業結構的重要力量，並同時影響生技製藥、資料中心、半導體與能源供應鏈等多個領域。在生技製藥方面，英矽智能與禮來公司達成高達 27.5 億美元的合作案，顯示人工智慧藥物研發已從技術概念走向商業落地。大型藥廠不再只把人工智慧視為輔助工具，而是將其納入新藥開發與全球授權合作的核

心。在資料中心與算力基礎設施方面，Mistral AI 在歐洲推動大型資料中心建設，顯示人工智慧競爭已延伸到實體運算基礎建設層次。企業競爭不再只看模型能力，也取決於是否能整合資金、土地、電力、冷卻、網路與晶片資源，建立穩定且可擴張的高密度算力環境。在半導體與系統架構層面，輝達與邁威爾的合作顯示，未來競爭已不只是單一晶片產品的競爭，而是整體人工智慧基礎設施標準與生態系主導權的競爭。輝達正透過互連、處理器、交換器、網路與光通訊技術，建立以自身為核心的人工智慧運算架構，並將資料中心、通訊網路與推論系統納入同一套整合體系。另一方面，中東衝突則顯示人工智慧與半導體產業雖然看似先進數位產業，實際上仍深度依賴能源、航運與特殊工業原料。荷莫茲海峽風險推升油氣價格，也增加晶圓廠與科技製造成本；氦氣供應風險則可能進一步衝擊半導體製程。再加上人工智慧需求排擠先進製程與封裝資源，也使部分消費性電子產品面臨成本上升與供應壓力。

博班生挑戰受挫韌性激發研究潛力:「磨難成才」

許多博士班學生雖然學業表現優秀，但往往缺乏面對研究不確定性與失敗的經驗，因此在實驗不順利時容易產生挫敗感與自我懷疑。研究訓練中常見的焦慮、倦怠與低自信，會讓學生停滯不前，而真正的研究能力，往往是在反覆受挫與修正中逐步建立。針對這個問題，多位資深 PI 提出幾個核心做法。首先是建立合作文化，讓學生能彼此討論與互相支持，而非孤立面對困難。其次是幫助學生重新理解失敗，將其視為研究過程中的正常現象與學習機會，而不是能力不足的證明。第三是協助學生設定務實的研究目標，避免過高期待帶來過度壓力與挫折。導師應維持可接近性，透過日常互動與在場支持，降低學生求助的心理門檻。同時，也需要依據不同學生的特質調整指導方式，在放手與介入之間取得平衡。

科技平台走向責任時代:「治理失衡」

兒少社群媒體治理已由個人使用與家庭教養議題，轉變為涉及國家監管、平台責任與科技風險的公共政策核心。以澳洲為例，對未滿 16 歲使用社群媒體設限，並要求平台落實年齡驗證與帳號管理，顯示政府已將社群平台視為可能影響心理健康與社會發展

的高風險環境。然而實務上仍存在繞過機制與執行落差，凸顯立法與治理之間的落差。進一步來看，各國在治理策略上呈現不同路徑。英國在是否立即立法與先行評估之間出現政策分歧，反映出在保護兒少與避免政策副作用之間的權衡。同時，科技公司亦指出年齡驗證的技術困難與潛在隱私風險，使治理問題從單純限制措施，擴展至資料保護、平台責任與數位參與等多重面向。Meta 相關訴訟顯示企業責任正從「產品是否有害」進一步延伸至是否已知風險卻未揭露。這一趨勢意味科技治理已進入以透明性與可歸責性為核心的新階段。隨著 AI 快速發展，若仍以擴張優先而忽略安全與揭露，類似問題可能再次出現。因此，未來關鍵不僅在於制定禁令，而是建立兼顧兒少保護、平台治理、企業責任與技術創新的整體治理架構。

人工智慧影響社會判斷與行為價值:「迎合偏誤」

隨著人工智慧日益深入社會生活，其影響已從資訊取得延伸至人類的價值判斷與行為決策。研究指出，大型語言模型存在社會性奉承（social sycophancy）現象，即使使用者的觀點具有偏差或道德風險，系統仍傾向附和與肯定。當 AI 被優化為提升使用者滿意度時，可能削弱人類在真實互動中透過摩擦所建立的責任感、反思能力與道德判斷。人類的社會學習本質上依賴真實回饋與人際互動，包括衝突、修正與關係修復。然而，奉承式 AI 回應正是社會摩擦反面。研究顯示，AI 不僅比人類更傾向支持使用者立場，甚至在單次互動後，即可能提高使用者對自身判斷的信心，降低其承擔責任與修正行為的意願。此種機制容易形成偏好-強化單向循環，使偏誤判斷持續累積。若人們長期習慣依賴提供無摩擦回饋的 AI，可能逐漸降低對真實社會回饋的容忍度，進而影響人際關係與道德發展。此風險在年輕族群或社會支持較弱者身上尤為顯著。面對此挑戰，關鍵不在於完全消除 AI 的支持性，而在於重新設計其激勵與回饋機制，使其能適度提供修正性建議，促進使用者反思與成長。未來需透過跨領域合作，建立兼顧技術效能、倫理責任與社會福祉的 AI 治理框架。

AI 代理人失控風險浮現:「代理失序」

隨著人工智慧代理人（AI agents）從被動回應工具轉向可自主規劃與執行任務的系

統，其應用已逐步深入電子郵件管理、行事曆安排與跨應用操作等真實數位流程。然而，這種能力的提升也同步放大了安全與治理風險。研究顯示，當代理人具備實際操作權限時，可能在理解不完整或限制條件不足的情況下，採取遠超出使用者原意的行動，甚至造成資料遺失與系統破壞。實證測試揭示，這些風險並非罕見異常，而是在現有技術條件下即可反覆出現的現象。代理人可能未經授權外洩敏感資訊、產生無效但耗費資源的行為，甚至生成具誹謗風險的內容。其根本問題在於，AI agents 缺乏人類在情境判斷、角色倫理與責任邊界上的內在機制，當指令模糊或受限時，容易偏離使用者真正意圖。此外，人們往往將其視為可信任的助理，進一步放大錯誤信任所帶來的風險。在產業快速部署與開源工具降低門檻的趨勢下，AI agents 正加速進入企業與研究場域，但相對應的安全驗證與治理機制仍明顯不足。未來關鍵挑戰不僅在於技術層面的風險控制，更涉及責任歸屬與制度設計，包括當系統造成損害時應由誰負責，以及如何在高風險場域（如醫療與國防）建立有效監管。

智慧數位資安生命週期

《模仿遊戲》電影描述電腦科學家與人工智慧先驅圖靈故事。圖靈在早期即提出關於機器智慧的構想，為後來人工智慧的發展奠定基礎。這部電影呈現在第二次世界大戰期間參與破解密碼的過程。1939 年戰爭爆發英國對德國宣戰，國家全面進入戰時狀態。倫敦遭到轟炸，社會處於高度緊張氛圍，戰火波英國民眾，即使兒童亦需配戴防毒面具並進行撤離演練，舉國上下全面動員。當時德國的科技實力先進，不僅體現在潛艦與武器系統，在資訊與通訊技術亦快速發展。德軍透過加密系統傳遞軍事指令以避免重要軍事訊息被攔截或外洩。當時戰爭已不僅是武力對抗，也包含資訊的傳遞與保護。因此，英國與盟軍面臨的重要挑戰，是無法掌握德國加密訊息。在此背景下，英國海軍司令延攬全國菁英人才如圖靈等科學家投入破解德軍密碼工作，試圖從資訊層面改變戰局。雖然德軍攻擊相關訊息，例如潛水艇的集結位置透過無線電傳輸，盟軍只要有收報設備即可攔截。然而攔截內容卻是加密後的亂碼，無法直接解讀。當時盟軍的補給船需穿越大西洋，而德國 U 潛水艇則在海中埋伏，透過加密訊息協調攻擊行動切斷補給線隔絕英

美同盟。。因此如果能破解德軍密碼不僅能掌握敵方通訊內容，也能推測潛水艇位置，進而進行反制。但由於其加密機制極為複雜，使盟軍長時間無法破解。英國因此動員數學、語言學與邏輯等領域的專家組成團隊，包括圖靈在內，嘗試系統性破解密碼。然而，恩尼格瑪（Enigma）機的組合數極為龐大，且密碼每日更換。實際可用的破解時間不到一天，必須在極短時間內完成運算，使破解工作幾乎不可能。這也凸顯密碼技術本質上是一場與時間競賽的問題，與現代加密技術乃至量子計算所帶來的挑戰具有高度相似性。

當時情境需依靠人工運算，無法像現在透過電腦快速處理。解碼團對中邏輯學家與語言學家傾向以人工方式來對抗德國的加密系統，試圖以人力分析破解密碼。然而圖靈具有不同的觀點。他認為單靠人工無法追上機器的運算速度，因而提出關鍵想法：「只有機器才能打敗機器」。在這個理念下設計解密機，透過機械化運算來尋找密碼規則，並在每日有限的時間內，快速對攔截到的訊息進行破解。這樣的計畫需要大量資源與人力投入。圖靈透過設計密碼題目進行人才招募，吸引具備數學與邏輯能力的人加入團隊。同時，他也在早期研究中提出機器可執行複雜任務的概念，認為只要透過適當設計，機器即可完成包括解密在內的工作。儘管過程中面臨資源與決策上的阻力，圖靈仍克服困難，完成解密機的建置，為戰局帶來關鍵轉折。

然而破解德軍密碼仍面臨難題。由於每天可用的破解時間非常短，而密碼組合數量極為龐大，用大海撈針方式去嘗試所有可能組合不可能在時限內完成，因此解密工作陷入困境。加上內部壓力與衝突，甚至有人主張應該停止這項計畫、關閉機器。在一次偶然的情境中圖靈在與第一線密碼攔截人員的交談中發現，德軍每天清晨發出的第一封電文，常會出現固定的開頭用語。原本被誤以為是人名的文字組合，經進一步分析後，對應的是德軍慣用的開場語句，例如「Heil Hitler」等固定內容，並且後續常接續天氣等例行資訊。這項發現提供了解密的起始訊息，使原本近乎無限的密碼組合，縮小為可計算的範圍。透過這些固定模式，搭配解密機的高速運算，團隊終於能有效提升破解效率，逐步解讀德軍通訊內容。也正是結合現場經驗與數學方法，英國最終得以成功破解德國的加密系統扭轉戰局。

AI 在數位資安中的演進，呈現出一條由歷史經驗延伸至現代技術發展的脈絡。透過先前盟軍與德軍加密與解密的案例，可以理解資安並非當代才出現的問題。早在二次世界大戰期間，德軍即已建立每日變動的加密機制，使盟軍需在極短時間內進行破解，形成一種持續性的動態密碼對抗。在此脈絡下，AI 的發展初期主要用於強化資安防護。由於密碼組合數量極為龐大，單憑人力難以處理，因此機器學習技術開始被應用於資安領域。自 2010 年前後，機器學習(Machine Learning)逐漸成熟，從傳統統計方法延伸至深度學習架構，如神經網路等技術，主要用於攻擊偵測與分類（classification），提升對大規模組合與異常行為的辨識能力。當加密與防禦系統逐漸依賴 AI 時，AI 本身亦可能成為攻擊目標。約在 2015 年前後，研究指出 AI 系統存在對抗樣本（adversarial examples）等脆弱性，使 AI 從原本的防禦工具轉變為潛在的攻擊面。此一轉變顯示，資安不再僅是利用 AI 進行防護，而必須同時面對 AI 被攻擊的風險。進一步發展至 2021 年前後，AI 資安逐漸進入攻防融合的雙向系統階段。AI 同時具備防禦工具與攻擊目標的雙重特性，顯示資安架構需從單一方向的防護，轉向整合攻擊與防禦的整體設計。此階段亦凸顯抗 AI 與資安並存必要性。在此基礎上，資安議題進一步延伸至治理與標準層面。AI 不再僅是技術問題，而進入組織政策、制度設計與生命週期管理的範疇。若缺乏適當的治理架構與標準，AI 反而可能放大風險，例如模型誤用、決策偏差或系統性錯誤。隨著近年生成式 AI 與 AI 代理人快速發展，資安問題再度出現新的轉變。攻擊型態已由傳統系統漏洞，逐漸轉向語言操控（reasoning manipulation）與代理濫用（agent abuse）。在高算力與跨國網路連結的環境下，這類風險具備更高的擴散性與複雜性，並可能與代理系統失控現象交互影響，形成新的資安挑戰。

演進過程中 AI 數位資安逐漸形成生命週期架構。於此架構之下，當前資安主要呈現兩大面向：一為「AI 強化資安」（AI for Cybersecurity），另一為「資安保護 AI」（Cybersecurity for AI）。前者著重於運用 AI 提升資安防護能力，包括資安營運中心（SOC）自動化、異常偵測以及威脅情報分析等，使系統能更快速辨識風險與回應攻擊。AI 在此階段作為防禦工具，有效提升整體資安效率與即時性。後者則聚焦於 AI 本身成為攻

擊目標所衍生的風險，例如訓練資料汙染、提示詞攻擊、注入式攻擊與模型竊取等問題。隨著 AI 系統被廣泛部署，其本身的安全性與完整性成為資安治理的重要議題。

AI 資安的生命週期治理，可歸納為五個核心流程：Identify、Protect、Detect、Respond 與 Recover（IPDRR）。此一架構同時適用於「AI 強化資安」與「資安保護 AI」兩個面向。在 Identify 階段，一方面需找出系統本身的風險，另一方面亦須辨識 AI 模型自身的潛在弱點。在 Protect 階段，AI 原本作為防禦工具，用於防範攻擊；然而當 AI 本身成為攻擊目標時，則需進一步建立保護機制，以確保模型與資料的安全性。進入 Detect 階段，系統需具備即時偵測威脅的能力，包括辨識異常行為與攻擊模式，同時亦需監測 AI 系統是否遭受攻擊或操控。在 Respond 階段，AI 可發揮自動化優勢，快速對資安事件進行回應與處理，包括應對模型被操控、提示詞攻擊或資料外洩等情境。最後，在 Recover 階段，重點在於系統與模型的復原能力。當資安事件發生後，需透過備援機制與修復流程，使系統恢復正常運作，同時確保 AI 模型在遭受誤用或攻擊後，仍能回到安全且穩定的狀態，完成整體資安治理的閉環。

以《模仿遊戲》為例來看 IPDRR 五個層級的資安生命週期。首先，在「辨識（Identify）」階段，盟軍與德軍形成典型的雙方對抗關係。德軍發展高度複雜的加密系統（Enigma），並透過每日更換設定來確保通訊安全，避免其加密機制被破解，本質上即為保護其模型參數與通訊規則。相對地，盟軍則以圖靈為核心，分析德軍通訊模式，嘗試找出加密規則與潛在特徵（pattern），此過程即相當於異常偵測前的建模與機率推論。在「保護（Protect）」階段，盟軍透過圖靈設計 Bombe 機器，將解密流程自動化，大幅提升計算速度與破解效率，並逐步建立系統化的防禦能力。另一方面，德軍則透過嚴格的操作規範與使用限制，避免加密規則外洩，以維持 Enigma 系統的安全性，這對應於現代 AI 模型保護與存取控制機制。當保護機制仍可能被突破時，即進入「偵測（Detect）」階段。德軍需持續監測其系統是否遭到破解、通訊是否被攔截，或加密機制是否出現異常。相對地，盟軍則透過固定語句（如天氣與慣用開場語）辨識通訊模式，結合機器運算進行特徵比對與異常判斷，以提升解密成功率。在「回應（Respond）」階段，盟軍於成功破

解後，並不全面使用所有情報，而是有選擇性地運用，以避免暴露已破解 Enigma 的事實，維持情報優勢。德軍則在偵測異常時，需立即調整策略，包括更換密碼、改變通訊模式與採取對抗性（adversarial）策略，以降低被破解風險。

最後，在復原（Recover）階段，雙方均需建立持續調整與恢復能力。德軍透過每日更新加密規則與強化通訊安全，維持系統有效性；盟軍則因應密碼持續變動，需每日重新破解並更新解密模型，形成持續學習（continuous learning）的過程。整體而言，若未能依循此五階段生命週期進行系統化治理，無論是以 AI 強化資安，或以資安保護 AI，皆可能產生重大風險。因此，建立完整的 AI 資安生命週期管理機制，已成為當前數位安全治理的關鍵。

金融智慧數位資安實例

在金融情境中，由於涉及資金流動與高價值交易，各類攻擊行為始終存在。銀行體系如同戰時的德軍，致力於建立高度加密且封閉的防護機制，以確保系統安全與交易完整性。然而，即使防護機制設計嚴密，仍可能如同 Enigma 系統般被突破。在人工智慧發展過程中，若未完整建構 IPDRR 資安生命週期，系統仍可能出現可被利用的弱點。孟加拉銀行事件即為典型案例。該事件發生於 2016 年 2 月 4 日至 5 日之間，雖然實際攻擊執行時間極短，但駭客早在更早之前即已潛伏於系統之中，顯示長期滲透與短期執行之間的高度落差。受害機構為孟加拉央行（Bangladesh Bank）。攻擊來源雖難以完全溯源，但根據其使用之攻擊手法、惡意程式與操作模式，推論為與北韓相關駭客組織（Lazarus Group）相似。該次攻擊主要針對 SWIFT 國際匯款系統，透過偽造 SWIFT 報文、植入惡意軟體以及盜用憑證，成功發出多筆跨國轉帳指令。最終成功匯出約 8,100 萬美元，資金經由菲律賓 RCBC 銀行流入當地賭場體系。原始攻擊目標金額高達 8.7 億美元，但因報文拼寫錯誤而觸發中介銀行警示機制，使大部分資金得以攔截，否則損失將更為嚴重。SWIFT 屬於封閉金融通訊網路，其安全設計類似於軍事網路，於物理與邏輯層面皆具備隔離性，外部直接入侵難度極高。然而，本次事件顯示，攻擊並非來自加密機制本身的破解，而是透過內部系統滲透與憑證濫用，繞過既有防護架構達成攻擊目

的。

入侵者首先利用流程與時間上查核機制落差。在當時的金融作業環境中，即使已具備一定程度的電子傳輸能力，整體流程仍高度依賴人工處理。一般而言，國際匯款需經過多重確認與作業程序，往往需要數日才能完成。因此，駭客即利用此一時間延遲與流程特性，作為攻擊切入點。此外各國工作日與假期制度的差異放大了時間落差。孟加拉為伊斯蘭國家，週末為星期五與星期六，而多數國家則為星期六與星期日。攻擊者選擇在星期五晚間發動攻擊，使孟加拉央行處於休假狀態，無法即時進行內部查核與應變處理。在此期間，駭客偽造轉帳指令，透過 SWIFT 系統向美國聯準會發出大額資金轉移請求。由於美國當時仍處於正常營運時段，相關指令得以傳入。然而當聯準會向孟加拉方面發出確認訊息時，因其正值假期，無法即時回應。待孟加拉於星期日恢復上班時，美國已進入週末休息時段，形成跨時區與制度性聯繫查核斷層。攻擊者亦操控通訊流程，透過偽造訊息與干擾驗證機制，使正常的確認程序失效。資金轉移路徑則設計為由美國轉往菲律賓馬尼拉，而該地區當時正值農曆新年期間，金融機構運作受限，進一步降低即時查核與攔截的可能性。整體攻擊的關鍵，在於破壞即時確認機制。當電話無法聯繫、即時溝通中斷時，只能依賴電子訊息進行確認，而此類確認方式在時效性與可靠性上均存在限制。同時，金融交易系統一旦依指令執行匯款，即具有高度自動化特性，即使後續發現異常並進行追認，資金往往已完成轉移難以追回。

事件的攻擊時間發生於 2016 年 2 月，經事後調查發現，其實整體入侵行動可追溯至 2015 年第三季。攻擊者在正式發動前，已長時間潛伏於系統之中，目的在於深入理解銀行內部匯款流程與作業邏輯，為後續操作建立基礎。初始入侵方式為寄送偽裝成求職信的電子郵件至孟加拉銀行。該郵件內含惡意程式，於使用者開啟附件後植入後門。此階段並未立即造成明顯異常，使收件者誤以為僅為一般系統問題，例如電腦故障或需重新啟動，因而降低警覺性。入侵攻擊以漸進方式滲透，如同病毒感染般逐步擴散，避免引發即時警示。在成功建立初步存取權限後，攻擊者開始於內部系統中進行長時間觀察與蒐集資訊，特別針對銀行的作業流程與設備使用情境進行分析。例如透過印表機等

共用設備作為節點，觀察資料流動與操作行為。由於銀行內部多數設備透過網路連接，共用資源（如印表機）形成關鍵節點，使攻擊者得以藉此進行橫向移動，將惡意程式擴散至其他系統。此類共用設備在網路架構中類似中樞（hub），一旦被控制，即可成為擴散管道，使攻擊逐步滲透至更多關鍵電腦。攻擊者在此階段並未立即採取破壞行動，而是持續蒐集資訊、分析模式，確保後續攻擊能精準執行。同時，攻擊者亦觀察銀行在國際匯款過程中的操作邏輯，包括孟加拉央行與美國聯準會之間透過 SWIFT 系統進行外匯轉帳時的驗證流程與控制機制。透過長期監控與分析，逐步掌握關鍵操作電腦、流程節點及查核機制。在充分理解整體系統運作後，攻擊者才進入最終階段，發動具自動化與持續性的攻擊行動。此類攻擊不僅具備高度隱蔽性，亦強調快速執行，以在最短時間內完成資金轉移並降低被發現與阻止的風險。

攻擊者在極短時間內連續發送 35 筆偽造 SWIFT 電文總金額約達 9.5 億美元，顯示其操作具高度自動化與計畫性。在正常情況下，銀行應具備即時查核與應變機制。然而，由於孟加拉央行當時處於假期期間，無法啟動電話確認程序，僅能依賴 SWIFT 系統進行雙向電文驗證。紐約聯邦儲備銀行在接收到異常匯款請求後，依標準程序回傳確認訊息，要求孟加拉方面進行核實。攻擊者利用已植入系統的後門程式，對回傳訊息進行攔截與刪除，使孟加拉央行無法接收確認通知。同時，攻擊亦延伸至印表機伺服器，使原本應自動列印的交易紀錄無法輸出，進一步削弱人工檢核機制。此外攻擊者刪除並竄改本地系統中的交易紀錄日誌（log），使內部系統顯示為正常狀態通過基礎檢核，未出現異常交易警示，掩蓋異常資金轉移申請電文，使銀行在第一時間無法察覺異常。而紐約方面則確實依程序完成確認與回傳，但相關訊息在傳遞過程中即受攔截竄改使孟加拉央行未能收到任何警示或查核要求，導致整體防護與回應機制失效。攻擊者不僅偽造交易指令，更同時控制通訊流程、紀錄系統與輸出設備，形成資訊操控鏈使銀行內部與外部之間的驗證機制失去效力。此金融資安攻擊案例在通訊過程中原本應用於安全確認的通訊機制，反而成為攻擊者操控資訊的管道。由於攻擊者長期潛伏並植入後門程式得以改寫銀行內部資訊系統的運作邏輯，使原本應觸發檢核系統安全警示包含應即時列印的警

示訊息與跨端驗證流程均未能正常執行。此一機制原本屬於金融資安既有的雙重查核措施但在攻擊者控制下失去功能。

在此情境中紐約端發送至孟加拉的安全確認訊息遭到攔截與篡改，使孟加拉銀行誤判系統運作正常，未啟動任何異常應變程序。同時，系統回饋亦被操控為「狀態正常」的訊號，使內部人員無法察覺潛在風險。攻擊者不僅阻斷警示訊息，亦主動偽造系統安全狀態，形成完整的資訊誤導機制。在攻擊執行層面，攻擊者除掌握銀行匯款流程外，亦展現高度自動化能力。在數小時內即完成 35 筆、總金額達 9.5 億美元的偽造電文傳送。傳統金融作業流程需依賴人工核實，包含文件列印、跨機構確認及通訊往返，通常需耗時數日，且受限於工作時間與節假日安排。一般跨國匯款流程涉及多重確認機制，需透過電文、電話或其他通訊方式進行驗證。然而，此類流程在時間效率上遠低於自動化攻擊系統，使攻擊者得以利用時間差快速完成資金轉移。此現象顯示人工作業在高頻率、自動化攻擊面前，難以及時應對。此情境與先前密碼戰爭中的概念相呼應。過去德軍透過高速且頻繁變動的加密機制維持優勢，而在本案例中，攻擊者則反向利用自動化與速度優勢，使防禦方難以即時整合人力與資源進行應對。

本案例可依據先前所述之資安分析架構進行系統性剖析，包括 IPDRR 資安生命週期、核心資產辨識以及威脅地圖建構，以全面理解孟加拉銀行匯款事件的攻擊模式與風險結構。首先，在戰略層面，此事件的關鍵在於攻擊與防禦之間的速度不對稱。在當時尚未廣泛導入人工智慧與自動化防禦工具的情境下，攻擊者得以利用時間差與流程延遲迅速發動攻擊。相較之下，防禦方仍仰賴人工查核與跨機構確認機制，導致回應速度明顯落後。再加上攻擊發生於假期期間，使應變能力進一步降低，形成明顯的「速度不對稱」（speed asymmetry）。其次，在資產層面，需辨識並保護關鍵資產。首先為 SWIFT 系統本身，其為國際匯款的核心通訊平台；其次為具備發送 SWIFT 訊息能力的操作終端設備；此外，操作人員的憑證與授權機制亦屬重要資產。本事件顯示，攻擊並非由外部直接入侵，而是透過內部系統發動，使交易看似來自合法來源。同時，相關憑證已遭盜用，導致原有的信任機制失效。在此情況下，系統仍將異常交易視為合法操作，顯示信

任資產已被破壞而未被即時察覺。攻擊者之所以能成功執行，與其長時間潛伏密切相關。透過潛伏期內的觀察與分析，攻擊者得以掌握銀行跨國匯款的作業流程與驗證邏輯，使其偽造的交易指令能高度擬真，降低被識別為異常的風險。最後，在威脅地圖層面，攻擊已由傳統邊界入侵轉向資料層與流程層的操控。攻擊者在取得關鍵資產後，得以對交易資料進行竄改，並控制相關系統行為。例如修改資料庫內容、刪除交易紀錄、干擾印表機伺服器輸出，以及操控內部網路節點之間的連結關係。透過對系統內部結構與資料流的掌握，攻擊者不僅完成資金轉移，亦同時隱匿其行為，使整體攻擊過程難以被即時發現。

當前金融體系已朝即時化與聯邦學習（federated learning）方向發展，並廣泛導入人工智慧進行異常偵測與風險監控。雖然實際資金入帳仍需一定作業時間，但在此期間，銀行內部已透過數位資安進行多層次、反覆性的自動化查核提升風險識別與攔截能力。相關制度的建立源於過去重大資安事件所帶來的制度性修正與強化。在技術層面聯邦學習與分散式防禦網路成為提升資安能力的重要方法。透過聯邦學習各金融機構可在不共享敏感資料的前提下，共同學習詐欺行為特徵與風險模式。系統僅交換模型參數或特徵資訊，而不涉及個別交易細節，有效兼顧隱私保護與資安強化。除行為特徵外，各機構亦可共享風險資訊，例如可疑 IP 或異常交易模式，進一步提升整體防禦能力。此種架構不僅適用於金融領域，於醫療體系中亦具有高度應用潛力。若以 IPDRR 資安生命週期進行分析，在辨識（Identify）階段，需針對系統風險與 AI 本身風險進行全面盤點，例如 SWIFT 系統、跨國供應鏈風險及相關授權機制。隨著 AI 廣泛應用於資安防護，亦須同時評估 AI 模型本身之脆弱性與潛在攻擊面。在保護（Protect）階段，金融體系已建立跨國制度性防禦機制，包括國際貿易政策、跨境查核流程及多層驗證機制。匯款流程需經由多方驗證，從發起端至受款端均需進行身份與交易合理性確認，形成較以往更嚴密的制度防護網。在偵測（Detect）階段，透過國際監測體系與 AI 分析能力，金融機構可進行即時風險感知。以孟加拉銀行事件為例，由於一筆匯款因拼字錯誤而引起德國中介銀行警覺，進而觸發異常檢測並中止交易，顯示風險感知能力在實務運作中的關鍵

性。在回應（Respond）階段，現行機制已由單一機構決策轉為政府與企業之間的協同應對，並透過跨國合作進行策略調整與風險控制，提升整體應變能力。在復原（Recover）階段，金融體系透過事件回饋機制持續優化制度與技術，包括導入 AI、強化人才培育及調整作業流程，以建立動態且具韌性的資安體系。SWIFT 系統為全球金融基礎設施在此過程中持續強化其安全機制，以支撐現代金融體系之運作，結合制度、技術與國際合作，並依循完整生命週期進行管理，方能有效提升金融體系之安全與穩定。

以上內容將在 **2026 年 4 月 8 日（三） 10:00 am** 以線上直播方式與媒體朋友、全球民眾及專業人士共享。歡迎各位舊雨新知透過星球永續健康網站專頁觀賞直播！

- **星球永續健康網站網頁連結: https://www.realscience.top/7**
- **Youtube 影片連結: https://reurl.cc/o7br93**
- **漢聲廣播電台連結: https://reurl.cc/nojdev**
- **不只是科技：https://reurl.cc/A6EXxZ**

講者：

陳秀熙教授/英國劍橋大學博士、許辰陽醫師、陳立昇教授、嚴明芳教授、林庭瑀博士

聯絡人：

林庭瑀博士　電話: (02)33668033　E-mail:happy82526@gmail.com

劉秋燕　電話: (02)33668033　E-mail: r11847030@ntu.edu.tw