115 年 2 月公務機密維護宣導

2026 年 AI 的負面衝擊與正面影響

根據 Google Cloud Security 團隊發表之 2026 年網路安全預測（Cybersecurity Forecast 2026），開宗明義即述明攻擊者正全面採用 AI 攻擊策略，AI 將成為網路攻擊之常態，提升速度、規模及精準度，報告說明 AI 的負面衝擊與正面影響。

負面衝擊

- 攻擊者全面採用 AI：AI 將成為網路攻擊的常態，提升速度、規模與精準度
- 提示詞注入（Prompt Injection）攻擊：透過操控 AI 模型，企圖繞過安全機制，執行攻擊者隱藏的指令
- AI 驅動社交工程：語音釣魚（Vishing）結合 AI 聲音複製模仿，攻擊更具欺騙性
- 影子代理（Shadow Agent）風險：員工私自部署 AI 代理，導致資料外洩與合規問題

正面影響

- AI Agent 安全挑戰：AI 代理將廣泛應用，身分與存取管理（IAM）為核心基礎，以動態存取控制，避免未授權或擴權

- 安全分析員角色轉變：AI 將自動化威脅分析，分析員專注於策略判斷

2026 年將成為攻擊者與防禦者共同邁入之人工智慧安全新時代，為維持網路韌性，組織必須優先部署主動且多層次之防禦策略，導入完善 AI 治理（AI Governance），以調整安全策略，動態因應威脅變化。

資料來源：節錄自「國家資通安全研究院 前瞻研究籌獲中心『2026 年全球資通安全威脅報告摘要分析』」