

附件、應用生成式 AI 及 AI Agent 工具廉政風險小叮嚀

風險主題	行為態樣	預防措施
機密洩漏	直接上傳未公開的標案、公文草稿或個資至生成式 AI。	去識別化 輸入資料內容不含機敏資訊，或優先使用內部建置的封閉式 AI 系統，並遵守資安規範。
著作權爭議	宣稱 AI 生成圖文為個人「原創」，或直接用於正式發布。	實質修改 將 AI 視為草稿工具，最終產出須經過人工檢視修改與確認。
虛假訊息	未經查證即引用 AI 提供之法條、數據或歷史研究結果。	二次查證 對 AI 提供的數據進行事實查核，確保公文書正確性。
科技詐騙	接獲長官或同仁語音/視訊要求變更撥款或帳號，立即執行。	多重驗證 透過第二種通訊方式（如手機、市話通話或面談）確認對方身分。
誠信揭露	在處理關鍵公務（如補助審查、研究內容撰寫）時使用 AI，卻刻意隱瞞。	誠實申報 遵循行政院及所屬機關（構）使用生成式 AI 參考指引，必要時註明內容來源。