負責任人工智慧治理政策

1. 目的

USI（以下稱「本公司」）致力於以負責任、安全、透明及符合倫理的方式，在其營運、產品、服務及支援性業務流程中開發、取得、部署及使用人工智慧（Artificial Intelligence，AI）。

本政策建立本公司對 AI 的高層級治理要求，以確保創新能在適當的風險管理、人員問責、資訊安全、隱私保護、法律與倫理合規，以及持續監督的基礎下推行。其目的在於促進值得信賴的 AI 成果，降低誤用或未受控部署的可能性，並為內部治理、稽核審查及利害關係人信任提供一致性的依據。

2. 適用範圍

本政策適用於所有董事、主管、高階主管、員工、承包商、臨時人員及其他經授權代表本公司使用、管理、監督、開發、採購、設定、整合或支援 AI 的人員。

本政策涵蓋：

- 內部自行開發之 AI
- 外部採購之 AI 服務
- 通用型 AI 工具
- 代理型 AI（Agent-based AI）能力
- 企業系統內嵌之 AI 功能

本政策適用於 AI 全生命週期，包括：

- 評估（Evaluation）
- 設計（Design）
- 採購（Acquisition）
- 導入（Implementation）

- 營運（Operation）
- 監控（Monitoring）
- 審查（Review）
- 退役（Retirement）

3. 治理與監督

AI 治理委員會（AI Governance Council）應作為公司 AI 活動之集中治理、監督與跨部門協調機制。其職責包括：

- 建立 AI 治理要求
- 審查重大 AI 風險
- 推動控制措施一致性落實
- 監督重大 AI 相關議題，包括高風險 AI 使用案例、事件及例外情況

AI 治理委員會應由相關控制與支援職能單位提供支持，包括治理、風險與合規（GRC）、資安架構、隱私法務、智慧財產權法務、企業架構，以及其他適當的業務或技術利害關係人。上述單位應提供專業意見，於必要時提出質疑與執行審查，並協助各自責任範圍內治理要求之落實。

重大 AI 風險、重大控制缺失、重大事件或高影響力 AI 使用案例，應依公司既有治理與報告機制向高階管理階層提報，包括資訊安全委員會（ISSC），必要時亦應提交董事會層級監督。本政策應依公司企業層級政策之正式治理程序，經董事會層級核准與支持。

USI 實施方式

USI 於 2026 年 6 月成立 AI 治理委員會，作為跨職能 AI 治理與監督機制。該委員會向資訊安全委員會（ISSC）報告，並每月定期召開會議，以檢視 USI 內部 AI 使用的最新狀況、追蹤重大治理發展，並審查需要額外監督、質疑或決策之高風險 AI 使用案例。

透過 ISSC，AI 治理委員會亦可定期向董事會層級報告 AI 相關治理與管理事項之狀況。相關會議治理紀錄及審查結果，均應依公司治理程序保存。

4. 負責任 AI 原則

4.1 可靠性與安全性（Reliability and Safety）

AI 系統於使用或部署前，應依其預期用途與風險程度進行適當程度的驗證與測試。

公司應採取合理措施，持續監控 AI 的效能、可靠性及輸出品質，包括識別異常行為、效能衰退、誤用或模型漂移（Model Drift）。如發現重大效能疑慮，應視情況採取審查、重新校準、再訓練或退役等措施。

如發現重大可靠性或安全性疑慮，公司應視情況及時採取更正、限制或停止使用等措施。

USI 實施方式

USI 對導入之 AI 解決方案建立審查流程，依使用案例之性質與風險，於導入前或導入期間提交 AI 治理委員會審查。AI 治理委員會的跨職能組成，旨在確保審查涵蓋可靠性、安全性、法規遵循、資訊安全、隱私、架構及法律等相關專業領域。

USI 亦運用 Microsoft Purview 等支援工具，強化第三方 AI 服務的合規監督，並使用監控能力持續追蹤 AI 使用及 AI 輸出之穩定性與可靠性。相關審查、監控活動及重大發現之紀錄，均應依公司治理程序保存。

4.2 資訊安全（Security）

AI 系統及相關資料應採用適當資訊安全控制措施保護，包括存取管理、日誌記錄、監控、安全組態設定，以及適用時的加密措施。

安全評估與測試活動應依 AI 系統的風險、關鍵性及暴露程度執行，並於相關情況下考量 AI 特有威脅。相關活動可依使用案例風險，包括弱點評估、滲透測試或等效安全測試，以及正式上線前與重大變更後之安全組態審查。

AI 相關安全事件與弱點，應透過公司既有的資安事件管理及韌性程序處理。

USI 實施方式

USI 透過治理審查、技術安全控制與持續監控的組合落實本要求。針對經核准的 AI 解決方案，AI 治理委員會及相關支援職能在執行審查時，會將安全要求納入考量，並特別關注存取控制、安全組態、資料保護及 AI 特有安全風險。

在適用情況下，USI 對 AI 相關環境採用身分識別與存取管理、最小權限存取、日誌記錄與監控，以及其他既有資訊安全控制。對於第三方或外部提供的 AI 服務，則使用支援工具與既有安全治理機制，強化對安全及合規風險之監督。

此外，AI 使用相關的安全事件、異常活動或已識別弱點，均透過公司既有的安全監控、事件管理與改善程序處理。相關審查、監控活動及重大安全發現之紀錄，均應依公司治理程序保存。

4.3 隱私保護（Privacy）

AI 使用應符合公司的數據分類、資料最小化、目的限制、保存期限及安全處理要求。

涉及個人資料、敏感資料或較高隱私影響之使用案例，應接受適當的隱私審查與治理。

隱私考量應整合至 AI 全生命週期，以支持合法、適度且透明的資訊處理。

USI 實施方式

USI 透過資料最小化檢查、限制於外部 AI 環境使用機密或個人資料、在適當情況下進行資料遮罩或去識別化、審查個人資料處理的合法依據與目的、執行保存及刪除控制，以及針對較高風險使用案例諮詢指定之隱私或資料保護職能，落實本要求。

USI 亦維持集團層級資料保護長（DPO）監督，以及一套隱私保護機制，以支持符合歐盟、亞洲及北美洲適用的隱私要求，包括於相關情況下執行資料保護影響評估（DPIA）等隱私審查與評估程序。

此外，公司每年提供隱私保護教育訓練，協助員工了解個人資料，以及在業務活動中，包括使用 AI 時，應如何保護個人資料。

4.4 公平性（Fairness）

在與使用案例相關的情況下，尤其針對較高風險 AI，公司應考量是否可能產生偏見、不公平待遇或重大差異化結果。

公司應每半年維持合理的監控、審查、質疑或公平性稽核機制，以識別偏見、不公平待遇或不適當輸出的跡象。針對較高風險使用案例，應依適用治理程序，於部署前、重大變更後及後續定期執行此類審查。

如發現重大公平性疑慮，應依其情境與風險，考量採取改善或額外控制措施。

USI 實施方式

USI 透過 AI 治理委員會對導入之 AI 解決方案進行跨職能審查，並特別關注預期用途、受影響的利害關係人、決策情境，以及使用案例是否可能造成偏見、不公平待遇或重大差異化結果。

AI 治理委員會的組成，有助於確保治理、隱私、法務、資訊安全、架構及業務等相關觀點均納入審查。對於較高風險或敏感使用案例，可能要求於允許使用或繼續使用前，執行額外審查、質疑或人工監督。

如發現公平性疑慮、偏見指標或重大審查發現，USI 可依情況要求額外防護措施、使用限制、改善行動或進一步評估。此外，USI 透過使用者訓練強調，使用者必須主動檢查 AI 產出，並對其使用、依賴或傳達的內容負責。相關審查、公平性評估、訓練重點及重大決策紀錄，均應依公司治理程序保存。

4.5 人工監督（Human Oversight）

人工監督程度應與使用案例的重要性與風險相符，尤其是 AI 可能重大影響決策、行動或外部結果時。人工監督應依風險程度與使用案例，透過適當控制模式確保，包括 Human-in-Command（人員保有最終決策權）、Human-on-the-Loop（人員持續監督）或 Human-in-the-Loop（須經人工核准）。

應記錄並保存必要的人工審查、監督行動、核准、介入、推翻決定、例外處理及重大決策，並符合公司的治理與紀錄保存要求。此類紀錄應能證明具實質意義的人員參與，並支持問責性、可追溯性、可稽核性，以及符合適用之法律、法規與治理義務。

除非依適用治理及法律要求明確授權，否則不得將 AI 視為敏感、高影響或其他受限制事項中的唯一決策者。

USI 實施方式

USI 透過治理機制，要求在 AI 的審查、核准及使用過程中保留具實質意義的人員參與，特別是針對較高風險或可能產生重大影響的使用案例。AI 治理委員會以及相關業務或流程負責人，協助確保 AI 不被視為敏感事項中的自主決策者，並確保適當的人類判斷持續納入決策流程。

實務上，包括針對較高風險使用案例設置核准檢查點、AI 產出在用於重大業務行動或對外溝通前進行人工審查，以及明確指定人員具有在必要時質疑、推翻、限制、暫停或停止 AI 支援活動之權限。

USI 亦透過訓練與治理要求強調，使用者必須主動審查 AI 產出、運用專業判斷，並對其所依賴、傳達或執行的決策或內容負責。必要的人工審查、監督行動、例外處理及重大決策紀錄，均應依公司治理程序保存。

4.6 透明性（Transparency）

公司應在符合情境、法律、合約義務或利害關係人期待的情況下，提升 AI 在相關業務流程、服務或產出中的透明度。

針對相關使用案例，公司應尋求確保適當使用者或審查人員，能以足以支持負責任使用、審查及有效人工監督的程度，理解下列資訊：

- AI 系統的預期用途
- 主要能力與限制
- 所使用的資料來源或資料類型（於適當且允許情況下）

- 相關風險與不確定性
- 在合理可行範圍內，AI 支援產出所依據的基礎或理由

當法律、合約、治理要求規定，或未揭露可能合理造成內容係由 AI 或由人類作者或審查者產生之誤解時，應對 AI 生成或實質由 AI 協助產生的內容進行適當標示、揭露或其他方式之識別。

USI 實施方式

USI 透過治理及溝通措施，使 AI 在相關業務活動中的角色可被看見、理解並獲得適當揭露。對已導入之 AI 解決方案，其預期用途、AI 在流程中的角色、關鍵限制及重大假設，均於治理流程中接受審查與記錄，使相關使用者及審查者了解 AI 如何被使用及其適用限制。

當 AI 產出或實質由 AI 協助產出的內容用於商務溝通、交付成果或其他相關情境時，若未揭露可能造成對其來源、性質或人員參與程度之誤解，USI 要求進行適當標示、揭露或說明。

USI 亦透過訓練及治理要求強調，使用者應了解 AI 產出內容的限制、於使用前驗證產出，並避免以誤導或不透明的方式呈現 AI 支援結果。相關揭露、治理審查及重大決策紀錄，均應依公司治理程序保存。

4.7 問責性（Accountability）

AI 相關治理、營運、審查、提報及改善責任，應透過公司的治理架構及職責分工明確指定。依適用情況，相關責任應涵蓋因 AI 使用所產生的錯誤、營運損害、合規違規、隱私影響、資安事件，以及法律或合約後果。

涉及 AI 的重大事件、控制失效或治理疑慮，應透過適當的管理程序進行調查。

AI 相關審查、評估、事件或稽核所產生的矯正行動與改善措施，應依適當方式追蹤至結案。

USI 實施方式

USI 透過為每個經核准的 AI 使用案例指定並記錄負責人、保存核准及審查紀錄、記錄例外與風險接受、定期向管理階層報告重大議題，以及保留足以支持內部審查、外部確信或適用稽核的證據，落實本要求。

AI 治理委員會負責監督治理一致性及重大例外事項；業務或流程負責人對其責任範圍內經核准 AI 的使用負責；資訊安全、隱私、法務、IT 與 GRC 等支援職能，則依各自職責負責審查、調查或控制措施。

此外，相關 AI 系統應保有可記錄關鍵操作及輸出或結果的日誌，使重大活動能依公司治理及安全程序，在必要時被追溯、審查及調查。

4.8 環境責任（Environmental Responsibility）

在相關且合理可行之情況下，公司應於 AI 能力的設計、選擇及運作過程中，考量資源效率、適度性及更廣泛的環境影響。

對於部署於公司自有或外部提供之資料中心的 AI 系統，公司亦必須考量能源效率、碳足跡及永續基礎設施作法，包括運算資源最佳化、冷卻效率，以及可行時使用再生能源。

USI 實施方式

USI 在選擇或營運重要 AI 能力時，會考量能源消耗、資源使用效率，以及相關情況下更廣泛的環境因素。對於內部託管的 AI 系統，USI 將環境考量納入資料中心營運，包括監控能源使用、最佳化運算工作負載，以及在適用情況下與企業永續目標對齊。

若使用第三方或外部提供的 AI，例如 Microsoft Copilot 服務，USI 亦將環境與永續相關因素納入供應商選擇及盡職調查。這包括評估供應商公開揭露的永續作法，例如 Microsoft 所公布之提升資料中心能源效率、增加再生能源使用，以及最佳化 AI 工作負載效率的方法。

Microsoft 的永續措施包括持續投資於節能資料中心營運、使用無碳電力，以及最佳化 AI 能源消耗與基礎設施效率的創新，這些措施支持了 USI 對環境負責任 AI 供應商之評估。請參考微軟官網說明 [https://learn.microsoft.com/en-us/industry/sustainability/advance-sustainability-ai]

此外，公司亦將供應商在資訊安全、機密性、透明度、授權，以及限制公司資料二次利用等方面的控制措施，納入供應商治理與審查，以確保符合 USI 的永續、風險管理及負責任 AI 目標。

5. 負責任使用（Responsible Use）

僅限經公司核准、授權或以其他方式允許之 AI 系統、模型、工具、服務或啟用功能，可用於公司業務。

本公司不會開發或部署任何屬於《歐盟人工智慧法案（EU AI Act）》所定義之禁止類別 AI 系統，包括操控型系統、利用弱勢族群、社會評分或未經授權之生物辨識監控。

AI 僅能用於合法業務目的，且應符合使用者被賦予之職責範圍及授權存取權限。

使用者仍須運用專業判斷，並應依情境以適當程度驗證 AI 產出，方可將其用於業務執行、內部決策或對外溝通。

此外，AI 系統應受明確使用邊界管理，包括授權範圍、預期用途及能力限制，以確保僅在核准情境內使用，且不超出其預期應用。

USI 實施方式

在 USI，本要求透過核准 AI 清單（Approved AI List）及相關治理機制來實施，要求員工僅能在已定義的使用情境及職務責任範圍內，為合法的業務目的使用經授權的 AI 解決方案。

所有新導入的 AI 解決方案均須依公司的 AI 治理程序，在導入前或導入過程中接受審查。審查過程中會明確定義並記錄各使用案例的預期用途、授權使用範圍及主要限制條件，藉此建立清楚的 AI 使用界線。

經核准的使用情境可確保使用者了解哪些類型的業務活動允許使用 AI，以及 AI 系統不得應用於超出其預期情境的範圍。

這些使用界線亦透過治理要求及使用者教育訓練進一步強化。相關要求強調，使用者必須遵循適用的輸入與輸出處理規範、避免使用未經核准的外掛程式或外部 AI 工具，且不得將 AI 用途擴展至核准目的之外，以防止非預期用途或任務範圍持續擴張（Mission Creep）。

此外，USI 亦採用核准機制、監控及日誌記錄等控制措施，以確保 AI 的使用始終維持在所定義的邊界內。所有使用者在依賴 AI 產出前，均須主動驗證 AI 生成內容，並對其利用 AI 所產生的內容、決策或溝通結果負責。

與 AI 使用相關之核准紀錄、已定義的使用界線、允許使用條件及重大例外事項，均應依公司治理程序妥善保存。

6. 禁止用途（Prohibited Uses）

禁止以 AI 的方式造成公司、客戶、員工或其他受保護資訊遭未經授權揭露、傳輸、保存或暴露。 除非已依適用要求完成特別審查、核准及治理，否則禁止於敏感、受限制或高影響領域中使用完全自動化的 AI 決策。

禁止未經授權使用 AI 執行跨系統操作、發起交易、進行大量處理，或繞過既有控制措施、核准程序及職責分離機制。

不得以 AI 從事非法、欺騙性、不道德、歧視性或其他受禁止的用途，包括不符合適用法律、合約義務或公司治理要求之使用。受禁止作法包括在適用法律或治理要求不允許的情況下，使用旨在實質扭曲人類行為的操控型 AI、非法社會評分，以及未經授權的即時生物辨識、監控或類似監視活動。

USI 實施方式

USI 透過治理限制、技術控制及針對不被允許之 AI 使用的管理升級機制，落實本要求。在適用情況下，未經核准之外部 AI 服務、外掛程式或整合功能，將透過既有的存取控制、租戶管理或網路管理措施加以限制；疑似違規或高風險例外情況，則須依公司治理程序進行管理審查與升級處理。

針對敏感或高影響 AI 使用案例，USI 實施強化之存取限制，包括角色型存取控制（RBAC）、最小權限原則，以及適用時的事前核准或多層級授權。此類使用案例亦可能接受額外監控、日誌記錄及定期審查。

USI 亦透過治理要求與使用者教育訓練強調，員工不得繞過核准控制措施、不得將 AI 用於未授權之業務目的，亦不得以可能造成重大資安、合規、隱私或倫理風險的方式使用 AI。重大違規、例外情況及相關管理措施之紀錄，均應依公司治理程序保存。

7. 風險管理（Risk Management）

公司應採用風險導向方法執行 AI 治理，並考量使用案例之預期用途、涉及資料之性質、自主化程度、影響範圍，以及對個人、營運或法規遵循義務的潛在影響。

當發生重大變更時，AI 使用案例應重新評估，包括用途、資料使用、整合範圍、使用者族群、決策影響或外部法律與法規條件之變更。

較高風險的 AI 使用案例應接受強化審查、治理關注及控制要求，包括適當時採取更嚴格的監督與升級處理。

已核准之 AI 使用案例應接受定期監控或審查，以確認風險假設、治理條件及控制有效性隨時間推移仍維持適當。

USI 實施方式

USI 透過治理程序，要求導入之 AI 解決方案及使用案例，依其性質、預期用途、資料敏感度、自主化程度，以及可能造成的業務或合規影響，進行審查、分類與追蹤。AI 治理委員會負責審查重大 AI 使用案例；較高風險情境會依情況於導入前或導入過程中接受強化審查與額外治理關注。

如發生用途、資料使用、整合範圍或決策影響等重大變更，相關使用案例應透過公司 AI 治理程序重新評估。USI 亦定期審查已核准 AI 使用案例，以持續評估演變中的風險、控制有效性及治理條件。使用案例審查、風險分類、重新評估及重大治理決策之紀錄，均應依公司治理程序保存。

8. 資料保護（Data Protection）

使用於 AI 之資料應依公司資訊分類及保護要求處理，並僅限於與相關目的相符、必要且經授權之範圍。

敏感、機密、個人或其他受限制資料，不得於外部或未經核准之 AI 環境中使用，除非已依適用之治理、資訊安全、隱私及法律要求明確允許。

AI 相關資料處理應於適當情況下，與公司的監控、日誌記錄、保存及事件回應能力整合，以支持問責性與可稽核性。

USI 實施方式

USI 透過資料治理措施，要求用於 AI 的資訊遵循公司資料分類與保護要求，包括限制在未經核准或外部 AI 環境中使用機密、個人或其他受限制資料。企業資料來源如需連接 AI 服務，應事先完成適當核准及治理審查；相關資料存取、傳輸或使用活動，則應於適用情況下接受日誌記錄與監控。

USI 亦將保存、刪除及存取控制措施納入整體資訊安全與資料保護架構，並透過供應商或合約控制，處理第三方 AI 供應商對公司資料的機密性、安全性及二次利用限制。相關核准、監控活動及重大資料保護發現之紀錄，均應依公司治理程序保存。

9. 培訓和意識（Training and Awareness）

公司應維持教育訓練與意識提升活動，以支持負責任、安全及合規的 AI 使用。相關訓練應於新進人員報到時或適用情況下於取得權限前提供，後續定期更新，並於治理要求、技術或風險條件發生重大變更後更新。

訓練內容可包括 AI 治理原則、資訊安全、隱私保護、輸出驗證、適當使用限制，以及對影子 AI（Shadow AI）與誤用風險的認知。

教育訓練與認知提升措施可依公司內不同角色、責任及 AI 使用性質調整。

USI 實施方式

USI 透過員工到職期間及每年提供 AI 相關訓練，並搭配針對相關使用者的角色別訓練、可接受提示詞與資料處理指引、常見 AI 失效模式認知，以及對使用較高風險或業務關鍵 AI 能力團隊的重點強化，落實本要求。

USI 亦定期透過電子報及類似溝通管道，向員工發布額外的 AI 相關教育訓練與認知內容，以強化重要治理、資訊安全、隱私及負責任使用要求。訓練對象可包括一般使用者、開發人員、系統管理員、核准人員、審查人員，或其他負有較高 AI 相關責任的角色。

10. 通報機制（Reporting Mechanism）

AI 相關疑慮，包括疑似誤用、控制弱點、重大錯誤、資安事件、隱私疑慮、倫理問題，或對具有重大影響之 AI 支援結果的申訴，均可透過公司既有的通報、倫理、資安、隱私或法規遵循管道提出。如適用程序設有申訴或救濟受理機制，公司應提供合理可及之管道，並依相關治理程序進行審查與回應。

已通報事項應透過公司既有治理與調查程序進行審查及處理，並視情況採取矯正或預防措施。

USI 實施方式

USI 透過既有的資安、隱私、法規遵循、倫理或其他治理疑慮之通報與升級管道落實本要求，這些管道亦可用於通報 AI 相關問題。當 AI 服務提供給外部客戶及第三方時，公司亦應提供 AI 專屬之外部通報管道。

已通報之疑慮，包括疑似誤用、異常輸出、控制弱點或重大事件，應依其性質與嚴重程度進行審查及分流，並視情況由相關職能單位參與。必要時，重大 AI 相關疑慮可透過公司治理程序升級處理，包括提交 AI 治理委員會審查。USI 亦依公司治理程序保存重大通報、調查、後續行動及改善結果之紀錄。

11. 治理指標（Governance Metrics）

公司得監控治理指標，以支持 AI 治理計畫之監督與持續改善。指標範例可包括經核准之 AI 系統或使用案例的數量與類型、已通報事件或疑慮，以及改善活動之狀態。

公司亦得追蹤訓練涵蓋率、治理審查活動、政策例外、與重大 AI 使用相關的環境或資源效率指標，或其他有助於呈現風險可視性及計畫成熟度的指標。

相關稽核觀察事項、審查結果或重複發生的問題，可作為改善優先順序及管理報告的依據。

USI 實施方式

USI 透過定期治理報告及指標審查，支持對 AI 相關管理事項之持續監督。AI 治理委員會每月審查活動之產出，可用於治理指標、趨勢分析、後續行動及管理報告。

相關指標可包括已核准 AI 使用案例之狀態、高風險案例審查、已通報事件或疑慮、訓練完成情況、改善進度，以及其他與 AI 治理有效性及成熟度相關的指標。透過資訊安全指導委員會（ISSC），相關治理報告亦可支持董事會層級定期掌握 AI 相關管理狀況。相關指標、報告及重大後續行動之紀錄，均應依公司治理程序保存。

12. 政策管理（Policy Management）

本政策應定期審查並視需要更新，以反映業務使用、技術、風險條件及適用法律或法規要求之變化。

AI 治理委員會負責維護本政策，並透過公司治理架構監督其實施。

13. 違規與例外（Violations and Exceptions）

違反本政策者，得依適用之公司規則、合約安排及當地法律要求，採取紀律處分或其他管理措施。

本政策之任何例外，均應依公司既有治理要求，完成適當之核准及風險接受程序。

附錄（Appendix）

本附錄以示例方式彙整禁止及高風險 AI 作法，包括操控、非法監控、歧視性使用、未經授權之自主行動，以及就業、金融、生產控制等敏感領域中可能需要強化治理審查的 AI 使用案例。

表 1：非法（禁止）AI 作法清單

下列 AI 作法違反適用法律與法規，或嚴重違反基本倫理原則，構成公司治理紅線，任何情況下均嚴格禁止開發、採購或使用。

類別	治理紅線（原則導向）	代表性範例（非完整列舉）
操控與欺騙	在個人不知情或違反其真實意願的情況下，使用 AI 操控其決策、行為或認知。	• 以潛意識或心理操控影響個人選擇 • 使用深偽內容冒充他人或誤導大眾
利用弱勢族群	針對弱勢族群，例如兒童、高齡者、身心障礙者或經濟弱勢者，利用其特定弱點。	• 以 AI 誘導高齡者進行不必要的購買或財務決策 • 蒐集或濫用未成年人的個人資料
非法監控與隱私侵害	在未取得法律授權或明確同意的情況下，部署 AI 進行大規模監控、追蹤或生物辨識。	• 未經授權的人臉辨識或生物特徵監控 • 使用 AI 推論員工情緒、心理狀態或行為傾向
社會評分與歧視	使用 AI 以具有社會歧視性的方式，對個人或群體進行分類、評分或差別待遇。	• 以社會信用評分影響服務或機會之取得 • 依種族、性別、年齡等受保護特徵給予歧視性待遇
公共安全與非法用途	將 AI 用於非法活動，或用於危害公共安全或國家安全之目的。	• 協助網路攻擊、資料竊取或詐欺 • 產生偽造證據或偽造法律、商業文件

表 2：高風險 AI 作法清單

高風險領域	風險說明（原則導向）	代表性範例（非完整列舉）
人力資源與人員管理	影響就業機會、職涯發展或勞資關係的 AI 系統。	• 招募篩選、履歷評估、面試評分 • 績效評估、升遷、薪酬或終止聘用建議
金融、信用與重大交易	對個人或組織產生重大經濟影響的 AI 系統。	• 信用核准、保險定價或理賠決策 • 大額資金移轉、投資或資產處分
關鍵基礎設施與生產控制	影響安全、營運持續性或系統穩定性的 AI 系統。	• 生產線控制、設備排程或品質判定 • 能源管理、供應鏈或物流控制
具有法律效果的自動化決策	產生法律效果或類似重大影響的自動化決策。	• 完全自動化的服務拒絕、處罰或限制 • 合約生成、法律或合規判定
代理型 AI（自主型 / Agent-based AI）	具備自主決策及跨系統執行能力的 AI 系統。	• AI 代理使用員工憑證跨企業系統執行任務 • AI 系統自主規劃及執行業務流程
敏感社會領域	應用於高度監管或公共利益領域的 AI 系統。	• 教育評量、考試評分或學術分發 • 醫療診斷、治療建議或健康風險評估