

人工智慧(AI)治理政策

2026/9/22

第一條：目的

為確保本公司於產品研發、製造、營運及服務過程中導入人工智慧技術之應用，能兼顧創新發展與營運風險控管，並符合相關法令規範與倫理要求，特制定本政策。
本政策旨在建立一套系統化之AI治理架構，以落實人工智慧應用之透明性、可控性及問責機制，避免因技術誤用或失效所衍生之營運、法規、資安或聲譽風險，並參酌金融監督管理委員會AI治理指引精神及國際標準（如ISO 42001與ISO 27001）予以設計；並透過AI技術之導入，提升企業營運效率、產品競爭力及數位轉型能力，以支持公司長期發展策略。

第二條：適用範圍

本政策適用於本公司及其子公司之所有部門、全體員工，以及與本公司業務相關之第三方，包括但不限於外包商、顧問、系統整合商及供應商。
凡於業務運作中使用人工智慧技術，包括自行開發或外部導入之系統（如機器學習模型、生成式 AI、影像辨識或 API 服務等），均應遵循本政策。

第三條：組織與治理架構

一、最終核決

董事長為公司 AI 治理之最高核決主管，並決策 AI 相關策略與政策及確保整體 AI 風險管理之有效性，已決策資源投入符合公司永續發展方向。

二、最高監督

總經理為公司 AI 治理之最高監督主管，並同時確認 AI 相關策略與政策，監督整體 AI 風險管理之有效性，已確保資源投入符合公司永續發展方向。

三、AI 治理推動小組：公司成立此任務型審查單位。

(一) 召集人或負責人：總經理或總經理指定之專案人員

(二) 主責成員：資訊處長（負責技術與資安）、法務處長（負責法律與合規）、稽核主管（負責風險評估）

(三) AI 治理推動小組職責：

1. 制定與修訂 AI 治理政策與管理制度
2. 建立 AI 風險分類與評估方法
3. 審議高風險 AI 應用案件
4. 監督模型生命週期管理與制度落實

(四) 執行與維運：每季向總經理或總經理指定之人員報告 AI 治理成效或重大事件即時通報。

四、AI 技術支援與安全小組：此小組是在資訊處下成立之執行單位。該執行小組主要負責評測工具安全性、建立企業內部安全 AI 平台、提供需確認之需求並附上相關證據給 AI 治理推動小組審查。

(一) AI 技術支援與安全小組職責：

1. AI 技術導入與平台建置
2. AI 模型管理與技術標準制定
3. AI 應用案例推動與跨部門整合
4. 建立 AI 平台落地機制
5. 支援風險評估與監控
6. AI 平台管理應符合資安 (ISO 27001) 及 AI 治理要求，並作為公司 AI 應用唯一正式平台。作為統一管控機制，其功能包括：
 - AI 模型部署與版本控管
 - API 管理與存取控制

- 使用紀錄與稽核追蹤 (Logs)
 - 資料安全與存取權限管理
 - 生成式 AI 安全控管
- (二) **AI 能力與訓練**：資訊處應與人力資源處一同建立分級 AI 能力培育制度，負責規劃與執行公司整體 AI 教育訓練，包含管理層、技術人員及一般員工，以強化 AI 使用能力與風險意識。並建立相應的人才教育體系，並建立 AI 能力評估與認證制度，以確保員工具備正確且安全之 AI 使用能力。
- (三) **執行與維運**：每季工作會議或重大事件即時通報 AI 治理推動小組。

第四條： AI 治理運作模式

本公司之 AI 治理運作模式，係針對人工智慧應用之導入、發展與管理，建立一致性之推動方式與控管流程，以確保 AI 應用具備可控性與持續性。在運作模式上，公司應從創新實驗至正式導入，建立分階段之管理機制，將 AI 應用區分為**概念驗證 (PoC)**、**試點實作 (Pilot)** 及**正式導入 (Production)** 等階段，並於各階段設置相應之審查與控制要求。前述各階段應明確訂定包括：

- 應達成之目標與驗證標準
- 風險評估與控管措施
- 審核與核決權限
- 系統上線與監控條件

透過上述運作模式，確保 AI 應用由研發至落地過程具備一致性、可追蹤性及有效之風險管理。

第五條： AI 治理原則

本公司於導入、開發、部署及使用人工智慧系統時，應秉持審慎負責原則，並遵循下列治理原則，以確保 AI 應用符合企業治理、風險控管及法規要求：

- 一、**合法合規原則**：人工智慧之應用應符合相關法令規範，包括個人資料保護、智慧財產權及資料使用法規，並應確保資料蒐集及使用具備合法依據，不得使用非法來源或未授權資料。
- 二、**人類最終決策原則**：人工智慧系統之功能應以輔助人類決策為目的，不得在未經適當授權或覆核之情況下，自動完成對公司營運、產品品質、客戶權益或法規遵循具重大影響之決策。對於高風險 AI 應用，應保留人員最終審核與決策之權限，以確保決策之合理性與責任歸屬。
- 三、**公平性與避免偏誤原則**：人工智慧模型應避免因訓練資料或演算法設計產生偏誤或歧視性結果。亦應避免因演算法偏誤導致對員工招募、績效考核或供應鏈管理中之人權侵害。相關單位應定期檢視其輸出結果之公平性與一致性，並於必要時進行調整或修正。
- 四、**資料最小化原則**：人工智慧系統於蒐集或使用資料時，應以完成特定任務所必需為限，不得過度蒐集、處理或使用資料。任何涉及公司機密、個人資料或客戶資料之資訊，應建立敏感資料使用與傳輸控管機制，以防止未授權外部 AI 使用造成資料外洩或濫用風險。
- 五、**透明性、可解釋性與問責原則**：人工智慧參與之業務流程，應能清楚辨識其使用範圍、功能限制及決策角色。對於重要決策或高風險應用，應具備適當之可解釋性，使相關人員得以理解其判斷依據。此外，應明確界定 AI 系統與人員之責任歸屬，確保在發生錯誤或爭議時，能夠進行責任追究與改善。
- 六、**可追溯性原則**：公司應建立完整之 AI 運作記錄與管理機制，以確保其決策過程具備可追蹤性與可查驗性。
- 七、**風險分級管理原則**：人工智慧系統應依其用途及影響範圍進行風險分級，並採取相應之管理與控制措施。風險等級越高，應適用越嚴格之審查、監控及覆核機制，以確保其運作符合公司風險承受度。
- 八、**安全性與穩定性原則**：人工智慧系統應經適當之測試、驗證及監控，以確保其運作之安全性與穩定性，並避免因系統錯誤、異常或外部攻擊對公司營運造成影響。

九、**負責任 AI 原則**：本公司之 AI 應用應遵循負責任人工智慧 (Responsible AI) 之理念，確保技術之開發與使用符合企業倫理、法規要求及社會責任。所有 AI 相關活動，應在兼顧技術創新與風險控管前提下進行，並避免對個人、客戶或社會造成不當影響。

第六條： AI 治理發展規劃

本公司應依營運需求、科技發展、資訊安全、及本政策所定之 AI 治理原則，分階段推動發展規劃，逐步提升治理成熟度與應用。發展規劃應分為**短期 (導入與盤點)**、**中期 (制度化)**、**長期 (治理成熟度優化)** 三階段推動並由資訊處負責統籌執行。風險等級應依其對營運影響、法規遵循、客戶權益及資訊安全之影響程度進行判定。

第七條： 風險分級、評估及模型生命週期管理

一、本公司應建立一致性之 AI 風險分級制度，依據 AI 應用對營運、法規遵循、資訊安全及公司聲譽之潛在影響程度，將其區分為高風險、中風險及低風險三類，並針對不同風險等級採取差異化之管理措施外，並納入嚴格審查與監控機制，且應強制執行事前風險評估、AI 治理委員會審查、人工作業覆核機制及定期重新驗證。

(一) 高風險

- **定義**：AI 的錯誤、偏差、失效、遭竄改或不當使用，可能直接或重大影響人身安全、產品品質、關鍵營運、客戶或員工重要權益、法規義務、機密資料或公司聲譽；或 AI 會直接觸發重大、不可逆或難以補救的行動。
- **案例**：產品放行或不合格判定、生產設備安全聯鎖、研發或製程參數自動控制、客戶資格或權益剝奪、法定申報內容自動產出並提交、直接執行高權限資安處置。

(二) 中風險

- **定義**：AI 主要提供建議、預測、分類、摘要或流程輔助，可能影響部門營運、客戶服務、財務或管理判斷，但在採取重大或不可逆行動前仍具備有效人工檢視、覆核及否決機制。
- **案例**：銷售預測、維修預警、品質異常提示、客服回覆建議、內部人力排程建議、供應商風險排序、內部文件分析、法務或採購文件初稿。

(三) 低風險

- **定義**：AI 僅用於一般效率提升、內容整理或低影響輔助，錯誤不會實質影響安全、產品品質、客戶或員工重要權益、法規義務、關鍵營運或公司聲譽，且結果容易被人員發現、修正或捨棄。
- **案例**：會議紀錄整理、公開資料摘要、一般翻譯與文字潤飾、簡報初稿、非正式腦力激盪、公開行銷素材草稿、程式碼格式整理或測試資料產生。

二、本公司所有 AI 系統均應納入完整之模型生命週期管理 (Machine Learning Lifecycle Management)。該生命週期應涵蓋自需求規劃、資料蒐集與處理、模型設計與開發、測試與驗證、正式部署、運作監控乃至模型退場或淘汰之各階段。

三、**內部控制三道防線機制**：本公司採行內部控制三道防線架構管理 AI 風險。

- (一) 第一道防線由使用單位負責，應確保 AI 之使用符合規範，並對其輸入資料與輸出結果之合理性負初步責任。
- (二) 第二道防線由法務、資訊主管及風險管理單位共同組成，負責 AI 制度建立、執行風險評估、控管措施設計與高風險案件審查，並對第一道防線 (含資訊處建置之 AI 平台與應用) 進行獨立之合規與資安監督。
- (三) 第三道防線由內部稽核單位負責，應定期檢視 AI 治理制度之執行情形與有效性，並提出改善建議。

第八條： AI 使用規範

- 一、AI 之使用應遵循以下要求，且使用人員應對 AI 使用行為及其結果負責：
 - (一) 不得將機密資料、客戶資料或未公開資訊輸入未經核准之平台
 - (二) 應評估其輸出內容之正確性、合理性及偏誤風險，並經適當人工審查
 - (三) 對於生成內容應採取適當控管措施，以防止誤用或誤導情形
 - (四) 生成式 AI 系統之使用應納入資安監控與日誌記錄機制
- 二、資料保護與匿名化：於 AI 模型開發及運作過程中，如涉及個人資料或公司機密資訊，應採取適當之匿名化、去識別化或遮蔽措施，以確保資料無法被直接識別。任何未經處理之敏感資料，不得直接用於模型訓練或輸入外部 AI 服務。
- 三、AI 內容標示與透明：本公司應確保由人工智慧生成或輔助產出之內容，於適當情境下具備可識別性。
- 四、對外提供或可能影響決策之 AI 生成內容，應依其性質採取適當之揭露或標示方式，包括但不限於註記、系統標示或其他技術措施 (如浮水印或識別標籤)，以避免使用者誤認為完全由人類產出。
- 五、本公司員工及合作夥伴於使用 AI 系統時，應遵循公司既定之使用政策與資訊安全規範，並應本於誠實信用原則審慎操作。在資料使用方面，任何人不得將涉及公司機密、客戶資料、技術文件或營業秘密之內容輸入至未經核准之外部 AI 系統，尤其是公開型生成式 AI 平台，以避免資訊外洩風險。
- 六、在生成式 AI 之應用上，其產出內容僅得作為輔助性參考，不得在未經人工審查之情形下直接用於對外正式文件、客戶回覆或重要決策。使用人員應對 AI 生成內容進行必要之查核與修正，以確保其正確性與合法性。
- 七、在委外或採購 AI 服務時，負責單位應於契約中明定服務範圍、資料使用權限、資安責任、保密義務及違約責任，並應評估供應商之風險控管能力，並要求供應商針對其訓練資料及模型產出不侵犯第三方智慧財產權做出擔保 (Warranty) 與賠償承諾，以降低公司因使用該 AI 服務而被控侵權的風險。若供應商無法提供相關擔保，應採取相應之風險對策或限縮使用範圍。此外，負責單位應爭取供應商優先選擇承諾使用再生能源或具備高效能 (Green AI) 的雲端服務供應商，以符合公司整體的範疇三 (Scope 3) 減碳與淨零路徑。

第九條： 異常處理與事件管理

當 AI 系統於運作過程中出現異常，如產生錯誤預測、重大偏誤、異常決策結果或疑似資料外洩情形時，相關單位應立即依內部通報機制通報主管與相關單位，包括資訊技術、資安及風險管理部門。公司應建立高風險 AI 系統之緊急關斷 (Kill Switch) 與人工接管程序，於必要時，應即時停止該 AI 系統之運作，以防止影響擴大。隨後應進行原因分析，釐清問題來源是否為資料品質不佳、模型設計缺陷、系統異常或人員操作不當所致。完成分析後，應擬定修正與改善措施，包括模型重訓、資料清理、控制強化或流程調整，並追蹤改善成效。如該事件已對公司營運、產品品質或客戶權益造成重大影響，或涉及法規風險，應提報 AI 推動小組並視情況進一步提報董事會。

第十條： AI 管理系統

本公司應建立制度化之 AI 管理系統，並逐步對齊 ISO/IEC 42001 標準，以確保 AI 治理具備一致性、可追溯性及持續改善能力。

AI 管理系統之建置應涵蓋政策管理、程序設計、風險評估、控制措施、績效監測及持續改善等要素，並透過「計畫、執行、查核、改善 (PDCA)」之循環機制持續精進。公司應妥善保存 AI 相關文件與紀錄，包括模型設計文件、測試報告、變更紀錄及決策依據，以支援內部管理與外部稽核需求。

第十一條： 資訊安全整合

本公司之 AI 系統應全面納入資訊安全管理系統 (ISMS) 之管理範疇，確保其符合 ISO 27001 標準所要求之資訊安全控制機制。AI 相關資產，包括資料集、模型、系統平台及 API 介面，均應列入資產盤點與風險評估範圍，並依其重要性採取適當之存取控制、加密保護及權限管理措施。如為外部導入或開源之 AI 模型，應要求提供軟體物料清單 (SBOM) 或針對開源漏洞進行定期掃描，以防止供應鏈攻擊。針對 AI 特有風

險，如訓練資料外洩、模型竄改、API 濫用或惡意提示攻擊 (Prompt Injection)，應建立相應之預防與偵測機制。

在監控方面，公司應建立分層監控制度：日常監控應以即時或每日方式進行，重點觀察 AI 輸出結果與系統行為；定期檢視至少每季進行一次，以評估模型效能與安全性；內部稽核則應每年至少實施一次，以確認制度之完整性與有效性。如發生 AI 相關資訊安全事件，應立即啟動資安事件應變程序，進行通報、控制及復原，並視情況通報 AI 治理委員會。

第十二條：監控、紀錄與稽查管理

- 一、AI 治理成熟度管理：本公司應建立 AI 治理成熟度評估機制，定期檢視 AI 應用、控制措施及管理制度之成熟程度。該評估結果應作為制度優化、資源投入及策略調整之依據，以持續提升公司 AI 治理能力。
- 二、本公司應建立完整之 AI 運作監控機制，持續追蹤各 AI 系統之運作狀況，包括其準確率、錯誤率、異常輸出情形及使用情境變化。依風險等級不同，應採取相應之監控強度。對於高風險 AI 系統，應建立即時監控與人工覆核機制；中風險系統應定期評估；低風險系統則得採抽樣檢查方式進行。所有 AI 運作相關紀錄，包括模型版本、輸入輸出紀錄、決策過程及異常事件，均應妥善保存，以確保日後可進行追溯分析與稽核查驗。
- 三、內部稽核單位應定期檢視上述機制之執行情形，並針對缺失提出改善建議。

第十三條：推動監督

為強化對 AI 風險之掌握，AI 治理推動小組應至少每季向總經理或總經理指定之人員報告相關資訊；如發生重大異常，包括 AI 系統失效、法規違反或對營運造成重大影響時，應即時通報並啟動應變措施。

第十四條：實施與修訂

本政策經董事長核定後施行，並應定期檢討其適當性，原則上至少每年檢討一次，以因應法規變化、技術發展及公司營運需求之調整。執行之細則辦法將於本政策通過後，另行建立，如有未盡事宜，悉依相關法令規定及公司其他內部制度辦理。