

定穎集團

負責任人工智慧政策與管理計畫

一、目的與適用範圍

定穎集團推動人工智慧應用，應兼顧營運需求、個人權益、資訊安全及環境影響。本政策訂定 AI 導入與使用之管理要求，作為各單位選用工具、評估應用及管理供應商之依據。

適用範圍包括自行開發、委託開發或採購之 AI 系統、模型與服務，以及既有設備或軟體新增之 AI 功能。

本計畫採分階段實施方式：導入前完成用途與風險審查，啟用時備妥必要管控，使用後追蹤結果並處理異常。各項措施之執行狀態，應以實際紀錄為準。

每項應用應於啟用前確認管理責任、檢查週期、異常處理及紀錄保存方式。涉及外部服務者，應確認公司與供應商之責任分工及可取得之管理資訊。必要管控尚未具備時，應限制使用範圍或暫緩啟用。

二、負責任人工智慧管理措施

1. 敏感 AI 功能之使用授權與存取限制

涉及人臉辨識、監控或其他可能重大影響個人權益之 AI 功能，應先確認業務必要性、資料使用依據及影響範圍，取得授權核准後始得啟用。

公司應限制可使用敏感功能之人員、用途及操作範圍，分別管理查詢、匯出與系統設定權限，並保留必要之操作紀錄。權限應依核定週期檢查，於職務異動或需求終止時調整或撤銷。

未經重新審查，不得將已核准功能擴大用於其他目的。

應留存資料：功能清冊、審查與核准文件、權限名單及檢查紀錄。

2. AI 生成內容與決策結果之明確標示

AI 生成內容提供他人使用時，應於文件、畫面或其他適當位置清楚標示其由 AI 產生。經人工審閱者，得另註審閱情形，但仍應保留 AI 生成之標示。

AI 所作成之判定、評分、排序或其他決策結果，應使接收者辨識其來源，並說明屬自動作成之結果，或供人員判斷使用之建議。

影響個人權益之結果，應併同提供人工覆核或申訴資訊。各應用負責人應於啟用前確認標示方式，並於使用後抽查其完整性。

應留存資料：標示規則、文件或介面範例及抽查結果。

3. 模型漂移與效能退化之偵測及修正

AI 模型應於部署前設定測試基準、檢查頻率與異常門檻。部署後應依用途檢查資料分布、輸出結果及效能變化，以辨識模型隨時間發生之漂移或退化。

例如，品質檢測應用可追蹤誤判與漏判情形；預測應用可比較預測值與實際結果。實際指標應依採用之應用設定。

發生異常時，應評估對產品品質、營運或人員之影響，採取人工覆核、限制使用、資料或模型修正等措施；必要時停止使用，待改善經驗證後再行恢復。

外部模型無法由公司直接調整時，應要求供應商處理，或採取版本更換、替代服務等方式降低影響。

應留存資料：測試基準、檢查結果、異常處理及改善驗證紀錄。

4. 已部署模型之公平性與偏誤定期評估

公司應為已部署之 AI 模型訂定公平性與偏誤評估週期，並於用途、資料或模型重大變更時追加評估。

評估應檢視資料代表性、結果差異及可能受到不利影響之對象。涉及招募、員工管理或身分辨識等應用時，應特別檢查是否產生不合理之差別待遇。

以設備或製程為主要對象之應用，亦應確認結果是否間接影響人員權益，並記錄公平性風險之判斷依據。涉及個人權益之評估，不得僅以整體準確率作為判斷標準。

發現偏誤時，應調整資料、模型或作業方式，並確認改善效果。

應留存資料：評估方法、週期、適用性判斷、結果及改善紀錄。

5. 降低 AI 資料中心與模型之生態足跡

公司應依 AI 之部署及使用方式，規劃降低能源、溫室氣體排放、水資源及設備資源負荷之措施。

採用外部服務時，應將 AI 資料中心與模型之環境資訊納入供應商評估，選擇符合任務需求且資源使用較有效率之方案，並透過供應商溝通或合作推動改善。

自行管理運算設備時，應檢查閒置資源、重複運算及設備使用效率，避免超出實際需求之資源配置。

各項改善應明定範圍、負責人與追蹤方式。環境成果應區分供應商整體表現及公司所使用服務之影響，並說明資料限制。

應留存資料：供應商評估、改善方案、合作紀錄及可取得之成效數據。

6. AI 決策或結果之異議受理與人工覆核

公司應於 AI 應用啟用前，備妥供使用者及受影響第三方提出異議之管道，並說明受理窗口、處理期限及覆核方式。適當之既有管道得予沿用，但應確認其可受理 AI 相關案件。

申訴應由具適當權責之人員檢視原始資料、AI 結果及其使用情境，不得僅由原模型再次運算後作成回覆。

公司應告知處理結果與理由。確認結果錯誤或造成不當影響時，應更正結果或採取補救措施；申訴人仍有異議者，應可申請再次覆核或升級處理。

案件需延長處理時，應說明原因及預計完成時間。不得因提出申訴而對申訴人採取不利措施。

應留存資料：受理程序、案件紀錄、人工覆核、結果回覆及補救追蹤。

7. AI 計畫與工具之永續影響量化

各 AI 專案應於評估階段辨識其可能影響之永續成果。具有永續改善目的之專案，應設定比較基準、衡量指標、統計範圍與資料來源，於實施後量化結果。

指標可依實際用途選用節電量、減排量、節水量、原料或廢棄物減量，以及職業安全風險暴露之變化。工具數量或使用人次，不應直接視為永續成果。

計算應考量產量、產品組合及其他同期措施，說明 AI 之貢獻與歸因限制。揭露淨環境效益時，應計入 AI 運算及相關設備增加之環境負荷。

預期、估算及實測成果應分別呈現，避免重複計算或將預測列為已實現效益。

應留存資料：基準、指標、原始數據、計算方法及成果說明。

8. 員工 AI 倫理使用與安全訓練

公司應向員工說明 AI 使用之核准要求、資料保護及禁止事項。相關使用人員應於工具啟用前，完成符合職務需要之倫理使用與安全訓練。

課程應涵蓋機密與個人資料保護、AI 生成內容與決策結果標示、錯誤與偏誤辨識、人工覆核、智慧財產、資安風險及通報方式。

應用管理者及技術人員應接受與管理責任相符之訓練。工具或用途重大變更時，應安排補充訓練，並依核定週期複訓及評估學習情形。

應留存資料：課程內容、適用對象、完成名單、訓練涵蓋率及學習評估。