合勤投資控股公司
負責任人工智慧政策

2026 年 9 月
第一版

合勤投資控股股份有限公司（以下稱「合勤投控」或「本公司」）致力於在產品研發、服務提供及營運過程中，負責任地使用與開發人工智慧（AI）技術，確保其符合法規要求、資安標準與永續發展目標，並兼顧客戶、員工與社會整體利益。

一. 適用範圍

本政策適用於：

- 本公司、其子公司及具有重大影響力之合資事業於其業務活動中所涉及的使用與開發人工智慧（AI）技術的行為。
- 公司內部 AI 工具之使用
- 嵌入於產品之 AI 功能
- 與第三方 AI 技術供應商之合作

二. 核心原則

1. 資料隱私與保護

在 AI 系統全生命週期包括: 設計、開發、部署、運作中，落實個人資料隱私的保護，並遵循適用的資料保護法規。

2. 網路資訊安全

在人工智慧系統的使用和/或開發的過程中，保護網路安全，並採取強健的資安措施，以防止網路攻擊，包括:

- 執行滲透測試、事件應變、加密與安全程式設計
- 採取防止有害或不適當內容的機制，如: 提示防護（prompt shield）
- 採用加密、身分驗證與存取控管

3. 公平性與避免偏見

AI 設計需確保公平，避免使用和/或開發中可能存在的偏見，包括:

- 建立資料治理機制，識別與減緩訓練資料中的偏見，避免對特定族群造成歧視性結果
- 定期執行 AI 偏見檢測與稽核
- 確保建立多元且具代表性的資料集

4. 人類監督（Human-in-the-Loop, HITL）

確保受 AI 影響的關鍵決策具備有效且可記錄的人類監督，並具備介入、覆寫或升級的權限與能力

- 監督形式包含：
  - 最終決策權（Human-in-command）
  - 監督模式（Human-on-the-loop）
  - 需人工核准（human-in-the-loop）
- 禁止 AI 在無人監督審查下，做出不可逆或高風險決策

5. 透明性與可解釋性

- 清楚揭露 AI 使用目的、情境、能力、限制、資料來源與風險
- 針對不同利害關係人提供其可理解的輸出原因說明
- 確保使用者知悉其正在與 AI 互動而非人類

6. 責任歸屬與治理機制

- 指定 AI 治理責任單位負責 AI 結果、風險決策與合規性的管理
- 建立風險通報，並制定調查與處理 AI 造成錯誤或損害事件的流程
- 確保 AI 決策與結果具可追溯性

7. 使用邊界與風險控管

- 定義 AI 系統的授權範圍、能力與限制的邊界，明確界定人工智慧可以做什麼和不能做什麼的界限，避免濫用、功能擴張或非預期應用
- 包括使用政策、防護機制、能力限制與升級管控機制
- 設置完善的防護機制，確保 AI 系統僅依其設計目的進行運作，避免被用於超出原先規劃的用途或情境

8. 環境永續

要求自身/第三方的 AI 資料中心/模型具備低環境足跡，以管理永續風險，包括:

- 衡量與降低 AI 運算所帶來的能源、水資源與碳排強度的影響
- 推動降低運算負載與能源浪費的最佳化技術
- 優先選擇高能源效率的基礎設施與設備，並鼓勵使用再生能源

9. 禁止性 AI 應用

本公司禁止使用或部署以下 AI 系統：

- 使用潛意識、刻意操控或欺騙性技術以實質影響個人行為的 AI 系統
- 利用年齡、身心障礙或特定社經弱勢，進行行為操控的 AI 系統
- 用於長期評估或分類個人可信度並造成不利待遇的社會評分系統
- 未經授權的即時遠端生物辨識系統（RBI）(即歐盟人工智慧法案禁止的系統）

三. 治理架構

本公司的「永續發展委員會」為於董事會下設置的功能性委員會，定期審查 AI 造成的風險、法規符合性及對 ESG 的影響，向董事會報告，並以董事會為最高決策單位。

四. 政策變更

本公司得因法令修正或營運需求不時更新本政策，最新版本將公布於本公司官方網站，並自公布時生效。

本政策經董事會核可後實施，修正時亦同。