

購案規範書

- 1. 品名：AI 系統安全評測方法與實測驗證委託案
- 2. 說明：
為強化 AI 評測技術與安全評測能力，針對生成式 AI 系統之安全性、可靠性及風險防護能力進行評測方法研析與實測驗證。透過 AI 紅隊測試技術建立安全評測項目及測試情境，並選定至少 3 個 AI 系統進行實際評測，比較不同 AI 系統於提示攻擊、越獄、不當內容生成及敏感資訊洩漏等情境下之評測結果，分析其安全風險與防護表現，據以提出 AI 安全評測指標、測試方法及評測流程建議，作為後續 AI 評測環境建置及評測技術發展之參考。
- 3. 功能 / 需求規格

序號	委託執行項目說明	驗收項目
1	AI 安全評測技術及產業應用分析	• AI 系統安全評測方法與實測驗證報告 1 式
	AI 安全風險與評測項目研析	
	AI 安全評測方法與測試情境設計	
2	AI 系統安全評測實測與紅隊驗證	
	AI 安全評測實測規劃	
	3 個 AI 系統紅隊實測	
	AI 安全評測結果分析	

4. 驗收交付及請款

- 驗收付款：廠商應於 115 年 12 月 18 日前交付「AI 系統安全評測方法與實測驗證報告」1 式。經驗收合格後進行請款。(驗收完成後依據工研院請款流程進行請款)

5. 訓練：☐有 ☒無。

6. 保固：☐有 ☒無

7. 服務：☒有 ☐無

- 廠商應具備完善之技術支援能力，並能即時提供系統建置、部署、測試及維運相關協助。
- 本規範書未盡事宜，得標廠商應依委託方需求及契約約定，配合辦理相關事項，並提供必要之技術支援、文件及系統調整，以確保本案順利完成。