本會於 APEC 品質基礎建設研討會
分享 AI 語言模型評測與認證制度之規劃與實踐

實驗室認證一處/高子桓

前言

因應全球對於人工智慧(Artificial Intelligence：簡稱 AI)的快速發展，人工智慧管理系統(AIMS)驗證機構及相關人工智慧產品測試實驗室之認證，已成為國際間認證組織關注的重點。國際標準化組織(ISO)已發布 ISO/IEC 42001:2024 人工智慧管理系統及 ISO/IEC 42006:2025 人工智慧管理系統驗證與稽核之標準，亞太認證合作組織(APAC)已將 AIMS 驗證機構之認證，納入 APAC 相互承認協議(MRA)之範疇。因應國際標準及認證發展之趨勢，本會評估將於 2026 年下半年開放人工智慧管理系統(AIMS)驗證機構之認證。

然而，對於 AI 產品測試實驗室認證，國際間尚處於發展的初期，現行國際認證組織關注的重點為已取得 ISO/IEC 17025 認證之實驗室，如運用 AI 工具在其校正/測試活動或實驗室管理作業，如何確保實驗室的公正性、能力及運作一致性可符合國際認證標準 ISO/IEC 17025，尚未提供 AI 產品測試實驗室之認證。我國數位發展部數位產業署為建立國內 AI 產品及系統評測體系，成立 AI 產品與系統評測中心(簡稱 AIEC)，並於 2024 年 3 月公告人工智慧（AI）產品與系統評測參考指引（草案），並陸續發布語言模型評測參考方法(草案)，以利國內實驗室及驗證機構，可提供產業在地化評測與第三方驗證服務。

因應國家政策的發展，本會於 2025 年度成立「人工智慧(AI)產品與系統評測實驗室認證研究工作小組」，並召開五次專家會議，初期聚焦於大型語言模型（LLM）測試實驗室認證之可行性及困難點，期望集結各界專家之意見，協助本會建構合理、可行且具前瞻性的 AI 測試實驗室認證制度。由於本會(TAF)已投入在 AI 語言模型評測實驗室認證之先期研究，故受邀參加亞太經濟合作組織(APEC)舉辦之國際研討會，分享我國建置 AI 產品與系統評測實驗室/驗證機構認證制度之構想及未來計畫，以下簡介本會參與 APEC 情況與本次會議主題介紹。

APEC 品質基礎建設(QI)研討會

亞太經濟合作組織(Asia-Pacific Economic Cooperation：簡稱 APEC)為促進亞太地區品質基礎建設(Quality Infrastructure, QI)之經驗交流，APEC 於 2026 年 1 月 20 日至 21 日舉辦「Conference on Best Practices of Quality

Infrastructure (QI) in APEC Economies」線上會議。本次會議由馬來西亞標準局(Department of Standards Malaysia)主辦，邀請各經濟體之標準、認證、計量及符合性評鑑等相關單位與會，進行跨經濟體之交流。

本次會議共規劃四大主題，其中 Theme 4「New and Emerging Technologies」聚焦人工智慧(Artificial Intelligence：簡稱 AI)、量子物理等新興領域所衍生之品質基礎建設需求。隨著語言模型(Language Model：簡稱 LM)快速發展，如何建立可信之 AI 評測與驗證機制，已成為各經濟體共同關注議題。本會受邀於 Theme 4 場次分享我國 AI 語言模型評測與認證制度推動情形。

TAF 分享主題及內容

本會高子桓經理，以「Language Model Certification for Artificial Intelligence – LM Evaluation and Accreditation Framework in Chinese Taipei」為題，介紹我國 AI 語言模型評測實驗室與認證制度。其分享要點如下:

認證面臨挑戰

國際標準化組織(ISO)已發布 ISO/IEC 42001:2024 人工智慧管理系統及 ISO/IEC 42006:2025 人工智慧管理系統驗證與稽核之標準，有關語言模型已廣泛應用於國內外相關產業，惟目前多數評測仍由開發者或使用者自行進行，缺乏第三方符合性評鑑架構、評測方法尚未形成一致標準，及結果可比性不足等問題。此外，語言模型輸出，易受模型版本、參數設定及系統架構影響，亦增加一致性測試與認證時之技術門檻。

我國推動 AI 評測現行規劃及做法

我國數位發展部目前推動 AI 評測制度，成立 AI 評測中心(AIEC)，參考 NIST AI RMF、ISO 24028 及 EU AI Act 等國際文件。建立 AI 產品與系統評測參考指引與參考方法，採自願性架構，依目前草案版之內容，第一階段將聚焦語言模型，設計十項評測指標，包括自動／半自動評測項目的公平性、準確性、可靠性、隱私、資安；與人工評測項目的安全性、可解釋性、彈性、透明性、當責性。為支持國家政策之發展方向，TAF 將於今(2026)年評估並規劃未來開放 AI 產品與系統測試實驗室認證(ISO/IEC 17025)，逐步建置我國 AI 產品與系統第三方符合性評鑑能量，以提升 AI 檢測產業之公正性、能力、與一致性。

會議回饋及交流

與會者就 QI 基礎建設如何幫助 AI 檢測驗證/認證產業發展提出經驗交流。指出透過 QI 體系之建立，可導入具公信力之第三方機制，以強化評測活動之公信力；提升結果之可再現性、可比性與可接受性，目標為增加我國 AI 評測體系之競爭力並提升跨組織甚至跨國接受度。與會者對我國嘗試以 QI 架構推動 AI 產品與

系統評測之作法表達關注，認為此一模式可作為其他經濟體發展 AI 符合性評鑑制度之參考。

未來展望

隨著大型語言模型快速發展，建立具公信力且可國際接軌之AI評測與認證機制，已逐步成為品質基礎建設之重要新興課題。惟目前國際間多數討論仍聚焦於運用 AI 技術輔助檢測驗證活動，AI 產品與系統本身之第三方評測與認證機制，整體而言仍處發展初期。

本會透過本次 APEC 品質基礎建設(QI)研討會，分享我國認證制度之規劃，展現我國以標準化為基礎推動 AI 品質基礎建設之方向。未來 TAF 將持續配合主管機關政策並積極與國際交流，逐步建構人工智慧管理系統(AIMS)驗證機構及相關人工智慧產品測試實驗室之認證服務。