

國家科學及技術委員會主管 115 年度單位預算評估報告

四、補助國研院繼續辦理「晶片驅動臺灣產業創新-精進臺灣可信任 AI 大語言模型推動計畫」，允宜依計畫執行實績，滾動式檢討並適度調高績效目標，俾利評估計畫成效

國科會 115 年度預算案補助國研院辦理「晶片驅動產業創新再升級計畫」編列 68 億 4,180 萬元，其中「晶片驅動臺灣產業創新-精進臺灣可信任 AI 大語言模型推動計畫」編列 1 億 1,880 萬元。經查：

(一)計畫概要

該計畫係 114 年度起執行之 4 年期計畫，銜接並精進 112 年度開始執行之「推動可信任生成式 AI 發展先期計畫」已建立之繁體中文版可信任 AI 對話引擎(TAIDE)，透過補助國研院建置資料蒐集平台，持續蒐集各面向專用領域及常用自然語言繁體中文語料，以官方政府資源為主、民間為輔，收納具臺灣主體性之文本資料，同時藉此精進模型開發技術，建立完備之臺灣自主大型語言模型、全民素養提升措施與組織學習標竿，並將 AI 模型應用於公部門施政或治理層面，藉以改善其作業流程與運作效能，期能促進臺灣之數位科技應用實力，115 年度兩項細部計畫之執行單位、計畫內容及預期關鍵成果詳表 1。

表 1 115 年度精進臺灣可信任 AI 大語言模型推動計畫概要表

細部計畫	執行單位	計畫內容	預期關鍵成果
資料蒐集及預處理	國研院科技政策研究與資訊中心	蒐集整理臺灣各面向之相關知識及應用開發所需之專用領域資料，維運並推廣資料蒐集平台，鼓勵民眾自主上傳其擁有著作權之資料。	完成資料蒐集平台之維運，持續蒐集繁體中文資料及至少 5 個應用主題之專用資料。
模型研發與應用推廣	國研院國家高速網路與計算中心	深化模型技術開發成果，掌握最新自然語言處理技術。發展多態模型，打造賦有價值觀與文化適應性之模型產出。打造「TAIDE-公部門公文生成模型」，以作為公	累計完成 5 套模型開發並提供特定用戶。

		部門以及多元領域模型導入應用之示範，期擴大可信任、本土化 AI 語言模型之推動基石。	
--	--	--	--

資料來源：115 年度政府科技發展計畫書「晶片驅動臺灣產業創新-精進臺灣可信任 AI 大語言模型推動計畫(2/4)」。

(二)截至 114 年 8 月基礎模型下載量已逾 22 萬次，雖 115 及 116 年度計畫下載量目標值已調升，惟仍未具挑戰性，允宜審酌執行實績設定具挑戰性之目標值，俾利評估計畫成效

該計畫原先設定開發之繁體中文基礎模型下載量目標值，114 至 117 年度分別為超過 1 萬、1.3 萬、1.5 萬及 2 萬人次，依 115 年度計畫書所載，下載量由「預期關鍵成果」改設為「自我挑戰目標」，下載量目標值調升為 115 及 116 年度分別達 1.5 萬及 2 萬人次，較原先設定目標值分別增加 2 千及 5 千人次，增幅分別為 15.38%及 33.33%。迄今 TAIDE 共有 7 模型版本，自 113 年 4 月開始陸續釋出，截至當年底累計下載量達 16 萬 6,967 次，至 114 年 8 月下載量為 22 萬 4,728 次(詳表 2)，已大幅超越計畫自我挑戰目標值，115 年度起調升之目標值仍未具挑戰性，允宜參酌執行實績，滾動式檢討並調升目標值，俾利評估計畫成效。

綜上，國科會 115 年度預算案補助國研院辦理發展推動本土語料模型開發，擴充資料來源與多樣性，持續精進 TAIDE 模型，並結合公部門應用推廣實現技術落地與強化數位治理。雖 115 及 116 年度計畫下載量自我挑戰目標值已調升，惟仍低於 TAIDE 模型實際下載量，允宜依執行實況滾動式檢討並設定具挑戰性之目標值，俾利評估計畫成效。

表 2 TAIDE 模型下載次數統計表

單位：次

模型名稱	釋出時間	模型差異	累計下載量	
			113 年 12 月底	114 年 8 月底
TAIDE-LX-7B	113/04/15	TAIDE-LX-7B-Chat 之 4 bit 量	2,275	3,041

模型名稱	釋出時間	模型差異	累計下載量	
			113 年 12 月底	114 年 8 月底
-Chat-4bit		化模型，以提供使用者之便利性為考量，可能會影響效能與更多不可預期之問題。		
TAIDE-LX-7B-Chat	113/04/15	以 TAIDE-LX-7B 為基礎，透過指令微調強化辦公室常用任務和多輪問答對話能力，適合聊天對話或任務協助之使用情境。	57,458	66,275
TAIDE-LX-7B	113/04/15	以 LLaMA2-7B 為基礎，僅使用繁體中文資料預訓練，適合針對模型進一步微調之使用情境。	10,596	14,306
Llama3-TAIDE-LX-8B-Chat-Alpha1-4bit	113/04/29	以 LLaMA3-8B 為基礎，使用繁體中文資料預，並透過指令微調強化辦公室常用任務和多輪問答對話能力，適合聊天對話或任務協助之使用情境。	3,783	5,165
Llama3-TAIDE-LX-8B-Chat-Alpha1	113/04/29	Llama3-TAIDE-LX-8B-Chat-Alpha1 之 4bit 量化模型，提供使用者之便利性，可能會影響效能與更多不可預期之問題。	92,855	115,105
Llama-3.1-TAIDE-LX-8B-Chat	114/02/11	相較於前期釋出之 Llama3-TAIDE-LX-8B-Chat-Alpha1，加強對於長文處理能力與繁體中文之解碼速度，以及模型上下文長度(Context Length)由 8K 增加至 131K。	-	19,482
Gemma-3-TAIDE-12b-Chat	114/08/22	相較於前期釋出之 Llama3.1-TAIDE-LX-8B-Chat，以 Gemma-3-12b 模型之進行訓練，並強化辦公室常用任務與多輪問答對話能力，適合聊天對話後任務協助之使用情境。	-	1,354
合計			166,967	224,728

資料來源：國科會提供。