

臺灣主權 AI 訓練語料庫

地方政府語料上架工作坊

一、辦理目的

為配合行政院主權 AI 相關政策推動，支援我國主權 AI 發展，本部已建置「臺灣主權 AI 訓練語料庫」，系統性徵集具臺灣文化特色及語意連貫之高品質語料，前已推動中央機關先行，逾 200 個政府機關投入，累計超過 3,000 筆資料集，語料規模超過 12 億詞元 (token)。

為持續擴大文本資料開放範疇及加速各地方政府享有著作財產權語料釋出，前已辦理地方政府語料徵集研商會議，請各地方政府就文本語料進行盤點與規劃釋出時程；為進一步強化地方協作及促進高品質訓練語料釋出，爰規劃辦理「語料上架工作坊」，採實作導向方式，手把手輔導地方政府熟悉平臺操作流程，完成語料資料集上架，攜手各地方政府共同支援我國 AI 模型之發展與應用。

二、主辦機關：數位發展部

三、執行單位：臺灣主權 AI 訓練語料庫輔導團隊

四、工作坊議程：

時間	工作坊內容
13:30-13:45 (15 分鐘)	報到
13:45-15:45 (120 分鐘)	語料資料集上架示範教學與實作
15:45-16:00 (15 分鐘)	回饋表單填寫

五、場次資訊：

場次		時間	地點
1	第一場 (高雄)	115 年 5 月 13 日(三) 13:30 至 16:00	高雄恆逸教育訓練中心 B 教室 (高雄市苓雅區新光路 38 號 4 樓之 1)
2	第二場 (臺北)	115 年 5 月 15 日(五) 13:30 至 16:00	國立臺灣師範大學和平校區 II 教育學院大樓四樓 403 教室 (臺北市大安區和平東路一段 129 號之 1)

六、報名須知：

- (一) 本工作坊邀請地方政府負責管理及上架 AI 訓練語料集之業務同仁參與，每一地方政府至多報名 5 名。
- (二) 報名方式：請各地方政府指派相關同仁，於報名截止日前至以下網址 (<https://gov.tw/eNd>) 完成報名資訊填寫。
- (三) 報名時間：即日起至 5 月 11 日 (星期一) 下午 5 時止。
- (四) 與會人員如需登錄公務人員終身學習時數 3 小時者，請於報名時提供身分證字號。
- (五) 錄取通知：報名後需經主辦機關審查，接獲錄取通知始得參加。
- (六) 候補須知：報名場次人數額滿後仍可登記候補，候補方式如下：
 1. 提供姓名、電子郵件、機關/單位名稱、職稱、聯絡電話/手機、預計報名場次、身分證字號(如需登錄公務人員終身學習時數者)等資訊，寄至 tsaitc@moda.gov.tw，候補成功者，將以信件與電話通知。
 2. 倘候補人數眾多，本部將視實際需求開放線上參與方式，提供無法到場之機關人員同步與會。
- (七) 主辦機關保留篩選受訓人員、調整授課場地與內容之權利。

七、與會須知：

- (1) 本次工作坊將於電腦教室舉辦，請各地方政府先行至「臺灣主權

AI 訓練語料庫」後臺 (<https://cms.taic.moda.gov.tw/>) 完成帳號申請。

- (2) 請各地方政府依據提交之語料盤點清單，備妥可編輯電子檔案 (PDF、ODT、JSON、CSV、XML)，於工作坊當日攜帶至現場，以利實際操作並完成資料集上架流程。
- (3) 與會人員將依公務人員請假規則規定給予公假，因故無法參加者，請提前告知以便安排候補人員。
- (4) 說明會場次時間、場地，如因故變更，將以電子郵件通知。
- (5) 為響應環保並落實節水行動，現場備有飲水機設備，請自備環保杯。
- (6) 機關需配合工作坊當日進行簽到程序，將全程錄影、錄音、拍照及滿意度問卷調查，相關蒐集資料僅作為主辦機關政策研析與專案管理使用。