

## 臺灣主權 AI 訓練語料庫十大文本資料類型

請各地方政府依下列文本資料類型進行內部電子資料盤點，徵集範圍包含但不限於清單所列內容，盡可能盤點及釋出可提供 AI 訓練之文本語料，支援我國主權 AI 發展。

	類型	範例說明	單位
1	政府出版品	縣市文化期刊 政府定期出版各類專業手冊（如： 防災手冊、長照資源地圖） 縣市年鑑	各地方政府
2	政府計畫/報告	縣市綜合發展計畫 年度施政計畫 出國考察報告	研考會、 各業務局處
3	自行/委託研究報告	文化資產修復調查 地方產業振興研究 環境影響評估報告	研考會、秘 書處、法制 局
4	政策與施政文件	政策文件如議會施政報告、各局處重 要政策白皮書	各地方政府
		政府公報如○○縣市政府公報	公報發行單 位
5	活動與宣傳文本	對外宣達文件，如新聞稿、專欄文 章、網路發布文章等	新聞局及相 關業務局處
		參與活動文件，如演講稿、致詞稿、 策展專刊、活動導覽手冊、各類節 慶活動文宣等文件	觀光局、文 化局及相關 業務局處
6	語言資源與詞彙 資料	本土語言推廣詞彙表 地方特色產業術語表	教育局、原 民局、客家

		地名淵源檢索資料	局或相關局處
7	教育與學習資源	訓練教材，如公務員培訓課程、政府內部學習手冊、職業訓練教材及相關題庫資料(含問答)	各地方政府
		科普教材，推廣民眾科學、技術、社會等知識	各業務局處
		教育教材，如中小學教材及相關題庫資料(含問答)、衛教資料、家庭教育等	教育部及相關機關
		語言學習教材，如華語文、臺語、客語、原住民族語、地方方言等學習教材	教育局及相關業務局處
8	臺灣在地文化	文化館藏資料，如地方志、館藏品資料	文化局或相關業務局處
		旅遊指南，包含美食、景點、節慶活動等介紹	觀光局或相關業務局處
		生物多樣性研究資料，如在地動植物調查報告等	農業局或相關業務局處
9	互動式溝通語料	客服對話資料(經去識別化)、常見問答(FAQ)等	各地方政府
10	其他可供 AI 訓練之文本	非屬上述類型但可作為 AI 訓練語料的文本資料	各地方政府