檔　　號：
保存年限：

數位發展部 函

地址：10066臺北市中正區延平南路143號
聯絡人：曾華伶
電話：02-23800392
電子郵件：lynntseng@moda.gov.tw

受文者：苗栗縣政府

發文日期： 中華民國115年1月12日
發文字號：數位資創決字第1153000040號
速別：普通件
密等及解密條件或保密期限：
附件：申請須知及使用規範(4058951_3000040_臺灣主權AI訓練語料庫申請須知及使用規範.pdf)

裝

主旨：有關「臺灣主權 AI 訓練語料庫」申請事宜，檢送申請須知及使用規範 1 份，請查照並轉知所屬單位。

訂

說明：

一、為推動臺灣主權AI發展，本部打造「臺灣主權AI訓練語料庫」（下稱語料庫），廣納高品質正體中文語料，支援AI模型訓練更貼近臺灣語言、文化與生活情境，促進AI模型具備更高的本土辨識力與語意理解能力。

線

二、語料庫目前已逾200個政府機關投入，上架累計超過3,000筆資料集，語料規模超過10億詞元(token)並持續擴充，收錄內容為各機關具臺灣文化特色之高品質資料集，涵蓋語言、文化、交通、教育、生物、地理環境等領域。

三、歡迎有AI模型訓練需求之機關(構)、公私法人、研究機構、學校、非法人團體或自然人申請使用，用臺灣的語料，打造理解臺灣的AI！

正本：行政院各部會行總處院（不含本部）、立法院、司法院、考試院、監察院、各直轄市政府、各縣(市)政府、中華民國資訊軟體協會、台北市電腦商業同業公會
副本：

電子公文交換章 2026/01/12 09:16:13

行政處 115/01/12
1150008472