115 年度臺灣主權 AI 語料貢獻獎勵活動簡章

數位發展部 115 年 7 月

壹、活動說明

一、依據「建構『臺灣主權 AI 訓練語料庫』行動計畫」辦理。

二、為持續擴充 AI 訓練語料資源，厚植我國 AI 發展及應用所需之語料基礎，推動各界投入語料整備與釋出，特辦理「115 年度臺灣主權 AI 訓練語料貢獻獎勵活動」（以下簡稱本活動），並設置「語料貢獻獎」，以表揚積極參與語料提供之機關及單位。

三、活動期間：本簡章發布日起至 115 年 10 月 31 日 23 時 59 分止。

貳、評獎方式

一、語料貢獻獎

（一）參獎對象

於「臺灣主權 AI 訓練語料庫」（以下簡稱語料庫）上架資料集之中央二級以上機關（構）（以下簡稱中央機關）、地方政府及民間為參獎對象。

（二）計算範疇

彙整計算參獎對象及其所屬機關（構）或所轄單位，已發布於語料庫前臺之資料集。

（三）分組方式

本活動依語料提供方屬性分為三組，各組獎項獨立計算及評獎。

組別	**參獎對象**
第一組：中央機關	中央二級以上機關（構），及其所屬機關（構）或所轄單位
第二組：地方政府	直轄市政府、縣（市）政府，及其所屬機關（構）或所轄單位
第三組：民間單位	公私法人、研究機構、學校、非法人團體或自然人

（四）評獎說明

1.計算方式

活動截止後，統一採用 OpenAI 發布 tokenizer（GPT-4/GPT-3.5 版本)計算參獎對象於語料庫前臺發布之資料集 token 數，並以最終加總 token 數為統計基準。

2.名次計算

(1) 分別計算三組參獎對象之 Token 數。

(2) 各組依 Token 數由高至低排序，取各組前三名。

(3) 同組總分相同者，並列同名次。

（五）獎勵方式

得獎機關（構）或單位由數位發展部予以公開表揚，獎勵方式如下：

1. 公開頒獎：於數位發展部公開活動或頒獎典禮頒發獎狀。

2. 成果公開：得獎名單及語料貢獻成果將公布於語料庫前臺，並得視活動規劃透過新聞發布等管道公開表揚。

參、其他事項

如有相關疑義請洽臺灣主權AI訓練語料庫維運管理單位：

（一）聯絡信箱：tsaitc@moda.gov.tw

（二）諮詢專線：0800-023-300（09:00-12:00 及 13:00-18:00）