

經濟部政府資料開放諮詢小組第 10 屆第 2 次會議紀錄

壹、時間：114 年 6 月 11 日（星期三）下午 2 時 00 分

貳、地點：經濟部第一會議室

參、主席：賴次長建信(林處長聰仁代理)

肆、出席人員：如出席人員簽名冊(如附件 1) 紀錄：蔡佳怡

伍、主席致詞：略

陸、報告事項：(報告單位：資訊處)

一、113 年度本部資料開放推動成果報告

決定：

(一)洽悉。

(二)請各單位持續配合本部及所屬機關(構)政府資料開放行動方案之策略，以及行政院政策，進行政府資料開放推動工作。

二、「經濟部政府資料開放諮詢小組」歷次會議結論追蹤報告

決定：

(一)洽悉。

(二)有關本部諮詢小組之歷次會議結論追蹤事項皆解除列管。

三、「114 年度政府資料開放獎勵活動簡章」說明

決定：

(一)洽悉。

(二)有關行政院 114 年度政府資料開放獎勵活動，請各單位配合下列事項：

1. 金質獎：

(1) 請各單位將所有資料集調整為 5 欄(含)以上，且確認資料集均達金標章。

(2) 有關白金標章資料集，請各單位就「人工檢核方式」進行確認。

(3) 如各單位有白金標章經平臺檢測未通過，而須提請申覆者，請於資訊處電子郵件通知之期限前，提交「資料標準辭庫對應建議表」。

(4) 如各單位接獲扣分項目調整通知，請儘速完成調整；未如期完成者，將依本部 114 年績效考核一併扣分。

2. 人氣獎：

各單位如有資料集活化應用情形，請於截止日前至平臺提報。

3. 文本語料獎：

俟數位發展部措施明確後，請資訊處於下半年諮詢小組會議說明或另案通知配合事項。

四、114 年度本部資料開放推動情形說明

決定：

(一)洽悉。

(二)有關所提 114 年上半年預計開放資料集共 29 項(如附件 2)同意開放，請提報機關於 114 年 6 月 30 日前依程序登載至政府資料開放平臺，並送資訊處進行審核發布作業。

(三)有關本部 114 年 6 月至 114 年 11 月資料開放工作重點及完成期限，請各單位配合下列事項：

1. 提送 114 年第 2 季及第 3 季累計已開放資料集檢核表(提送時間依資訊處電子郵件通知)。

2. 辦理 114 年下半年開放資料集提報作業(114 年 9 月 15 日完成)。

柒、討論事項：(報告單位：資訊處)

一、本部「能源管理」高應用價值資料主題-新增資料集討論

決議：

有關本部「能源管理」高應用價值主題新增 4 項資料集討論(資料集明細及審查結果如附件 3)，審查結果如下：

(一)3 項同意納入：

1. 有關臺中市政府經濟發展局提請新增「臺中市加氣站清冊」及「臺中市加油站資料」等 2 項資料集，考量資料集內容符合「能源管理」主題-「能源資源」子類別定義，爰同意該 2 項資料新增納入。
2. 有關臺中市政府環境保護局提請新增「臺中市電動汽機車充電站相關資訊」1 項資料集，考量資料集內容符合「能源管理」主題-「電力供需」子類別定義，爰同意該項資料新增納入。

(二)1 項不同意納入：

有關臺中市政府水利局提請新增「臺中市政府水利局水資中心回收水資源免費供市民取用量一覽表」1 項資料集，考量資料集內容與「能源管理」主題之定義及範疇不符，爰不同意納入。

(三)請資訊處將資料集審查結果登錄至政府資料開放平臺。

捌、臨時動議：

案由：陳委員孟敏針對「文本語料開放授權法律風險」提供之建議建議可進一步洽詢朱委員瑞陽針對「開放文本語料」相關之著作財產權問題進行研商，包含語料使用於 AI 訓練時的授權風險、資料開放適法性，及實務上可行的開放與授權模式，以利後續各機關在開放與應用文本語料時有明確依循與法遵依據，降低潛在法律爭議，並兼顧推動本土 AI 語言模型發展所需語料資源的穩定取得。

決定：

感謝委員提供之寶貴建議，將作為本部後續推動政府開放資料作業之重要參考。

玖、散會。(下午 3 時 00 分)

拾、附錄(與會委員發言紀要)

附錄、與會委員發言紀要

一、報告事項：

(一)朱委員瑞陽：

1. 為推動 AI 語言模型訓練所需的繁體中文語料，建議各機關盤點並考慮開放文本資料，採用創用 CC 授權方式，明確授權條件（如標示、非商業、相同方式分享等），以利後續安全使用。
2. 雖然 AI 模型的訓練階段本身不構成著作權侵害，但模型產出內容仍有潛在著作權風險，故語料的授權與來源應特別審慎。
3. 鼓勵機關不因著作權疑慮而完全不開放，而是思考如何在合法、安全的框架下促進資料可用性。目前多數 AI 模型的訓練語料以簡體中文為主，臺灣本地語言風格、用詞語感無法反映。
4. 若不積極開放繁體中文語料，AI 工具將越來越偏向簡體中文腔調與語感。

(二)葉委員慶隆：

1. 目前臺灣 AI 模型訓練所用的文本語料，多數並非純粹來自本土原生資料，而是從國外或非本土來源蒐集，再經後續清洗和微調使用，因此在模型輸出時常出現帶有對岸腔調的用詞與語感，影響本土語言特色。
2. 部分文章或信件開頭會出現過度形式化甚至不自然的尊稱，建議經濟部與數位發展部在建置與開放文本語料庫時，需加強語料篩選與詞彙過濾，確保資料符合臺灣本地語言特色和文化，同時在資料授權上維持透明與合規，確保未來 AI 模型訓練能彰顯臺灣語言主權，避免國外語料主導本地 AI 發展。
3. 所有文本資料在進入訓練前，應經過嚴格的清洗與整理程序，避免不適合的語言元素混入，提升模型質量與在地化表現。這是確保臺灣 AI 發展自主性與競爭力的重要基礎。

(三)包委員蒼龍：

1. 在協助教育部推動大學開放資料集的過程中，我們針對資料集內包含外部連結的部分進行討論。考量到連結有效性問題及資安風險，擔心部分連結可能指向惡意網站，影響使用者安全，因此初期決定不納入可點擊連結的內容。後來為兼顧需求，改採將相關內容轉成 PDF 檔案形式提供，避免連結失效問題，同時降低惡意網址帶來的安全風險，確保使用者在查閱時的電腦安全與資料完整性。
2. 鼓勵機關持續提升資料集品質並朝白金標章努力。

(四)陳委員孟敏：

OAS (OpenAPI Specification) 認證的要求符合國際趨勢，值得嘉許；但也確實對結構化、標準化的資料接口有較高要求。在規範的過程中，需注意避免為了規避這些較高要求，而讓開放資料走進只以 PDF、CSV 格式開放的局面；以 API 的介接，能有效達到產業化的最佳效益。也是政府資料開放平臺驗證、金質獎評選，對資料後續應用上的最終期待。

二、討論事項：

(一)包委員蒼龍：

針對資料集需具備五個有效欄位的要求，實務上部分資料欄位如說明文字或機關代碼可能不符合標準，導致欄位數量不足或欄位內容不適用。建議可透過拆解複合欄位(如地址拆為縣市、鄉鎮等)增加欄位數量，並適度放寬欄位類型限制，允許有助於資料理解的文字欄位納入。同時，應優先保留關鍵欄位，確保資料品質與應用價值，並考慮對欄位數量要求提供彈性調整，以提升資料集的實用性與開放效益。

(二)朱委員瑞陽：

針對再生能源發展相關資料，目前未納入水資源相關資料集部分，建議可邀請水利署與台灣自來水公司共同研商，評估是否

可另行訂定一項具高應用價值之新主題，整合再生能源與水資源管理資料，提升資料開放效益，並促進跨領域資料應用與政策決策支持。