

115 年度文化部資料開放諮詢小組會議

會議紀錄

壹、時間：115 年 8 月 3 日(星期一)下午 2 時

貳、地點：本部徐州路辦公室第 1 會議室

參、主席：徐委員宜君

紀錄：蔡大慈

肆、出席委員：莊委員舜清、連委員丁幼(李長龍 代)、陳委員宜靜、陳委員
薇如、孫委員玉達、林委員美婉、朱委員斌好、林委員玉凡、
劉委員嘉凱

伍、請假委員：戴委員豪君

陸、列席單位：本部資訊處、文化資源司、國立臺灣歷史博物館

柒、主席致詞：略

捌、報告事項：

一、本部資料開放推動情形(報告單位：本部資訊處)：略。

二、資料開放平台轉移作業(報告單位：本部資訊處)：略。

三、「115 年度政府資料開放中央部會第 1 次工作小組會議」重點報告(報告單位：本部資訊處)：略。

四、行政院修訂要點及獎勵措施(報告單位：本部資訊處)：略。

五、文化部所屬博物館館藏數位資料補實、開放 及利用計畫追蹤辦理情形(報告單位：本部文化資源司)：略。

六、權利盤點執行現況報告(報告單位：國立臺灣歷史博物館)：略。

玖、發言要旨(依首次發言順序)

一、林委員玉凡：

(一)針對數發部 AI-Ready Data 檢測作業之標準，若貴部檢視資料集情形，發現其準則實有窒礙難行之處，建議向數發部反映。例如資料欄位之規範，若盤點之資料集原始資料僅 4 個欄位，是否有必要為達標準而增加欄位，此狀況需考量向數發部提出建議。

(二)關於文本語料要開放到產業共享的時候，就資料流通方面，建議

洽數發部了解詳細內容。因為最近很多案例裡面，資料拿去訓練之後還是有些 bias 沒有克服，所以建議要了解後續資料的擴散狀況是如何。

- (三) 館藏盤點隨著新藏品進入，工作量勢必增加，目前國內外許多博物館利用 AI 協助藏品入檔，並建議考慮未來新藏品進入後建立對應的 LMV 型錄，於權利不清楚時可以參照使用。以內部使用來說，著作權很重要，適合用在初盤時期，但因為若依賴藏品原有的文件可能需大量的行動考察，此時在初期有其他可參照的外部資料時，可以有效加速盤點流程。

二、劉委員嘉凱：

- (一) 我認為數發部推 AI-Ready 是階段性推動，資料品質的下一步應該是資料的交互使用。例如各博物館有不同的 CMS，各館格式也不一定一樣，管理的邏輯也不一樣，那如何達到民眾需要時，可以用一個問答去把所有館的相關資料脈絡都找出來？那 AI-Ready 的下一個階段就是走到 MCP，貴部也許可以考慮先往前規劃，例如變成採用 MCP 方式主動對外提供這種的 AI 資料集。
- (二) 對於開放文本語料釋出建議保持謹慎態度，考量文本是整個歷史的累積結果，數位化之後提供出去被如何使用或拿去訓練 AI 模型的話無法看見。
- (三) 認同臺史博在進行權利盤點時提出的初盤方式。在緊急醫療裡有所謂檢傷制度，對照過來的話，即是分級制度。建議依盤點目的例如建檔用還是開放資料用或其他，針對不同用途別從 14 種權利中選取對應要盤的權利，建立快速分流方式，對於沒把握的部分再進行 2 次盤點。

三、朱委員斌好：

- (一) 關於民眾反應部分，好像比較多比例是因為網頁網站改版原因，那若民眾有反應貴部會知道，但如果沒有反應到是不是可能造成

部內不知情。建議改版前內部可以先作測試，確保不會有過多影響。

(二) 另外關於演唱會活動、涉及文化部或地方文化局的相關補助活動，是不是可以考量在合約中要求廠商提供相關資料，以達到豐富開放資料內容並且滿足民眾需求的可能性。

壹拾、 會議決議：

- 一、有關前次會議決議事項第一案數發部預定新增高應用價值「多元文化」專區，考量尚待原民會後續推動狀況，本案持續追蹤。
- 二、數發部推動 AI-Ready Data 檢測作業之標準請再檢視是否有窒礙難行之處，並適當向數發部反應。
- 三、因應本部自建之資料開放平臺下架作業，本部綱要計畫伍、一、「建立文化資料開放服務網整合平臺，並負責系統建置與維護、資料介接整合等」之文字比照刪除。
- 四、本部開放文本語料提報共 131 項資料集，並未全數通過審核，請洽數發部了解相關原因及審核標準。
- 五、配合行政院修正「政府資料開放諮詢小組設置要點」，本部諮詢小組會議調整為每年召開一次，必要時可加開會議。117 年由本部主辦之總統盃黑客松應及早啟動，若有需討論事項可於諮詢小組會議中提出。
- 六、有關「文化部所屬博物館館藏數位資料補實、開放及利用計畫追蹤辦理情形」，請文化資源司邀請工藝中心或館藏性質較臺史博不同的博物館，於下次諮詢小組會議分享盤點相關經驗以進行交流。

壹拾壹、散會：下午 3 時 30 分