

國立故宮博物院

113 年度政府資料開放諮詢小組第 1 次會議

會議紀錄

壹、會議時間：113 年 8 月 15 日（星期四）上午 10 時 30 分

貳、會議地點：線上會議

參、主持人：黃委員兼召集人永泰

記錄：游閏雅

肆、出席人員：顧委員振豪、林委員開世、廖委員鳳玓、林委員誠夏、張委員瑜倩、本院綜合規劃處張處長嬋娟、器物處吳處長曉筠、書畫文獻處何處長炎泉、行銷業務物王處長耀鋒、展示服務處王處長芸芬、登錄保存處林處長國平、南院處彭處長子程、主計室林主任美瑤、數位資訊室謝主任俊科、陳專門委員中禹、林科長育生、游助理研究員閏雅

伍、主席致詞：(略)

陸、報告事項

一、本院資料開放業務（詳見會議簡報）

二、112 年第 2 次會議決議執行情形（詳見會議簡報）

柒、討論事項

一、議題討論：因應國際趨勢發展及我國資料應用需求，擬進行 AI-ready data 前置作業，並規劃未來可應用之方向。

說明：為配合現行 AI 趨勢並擬製作符合大眾與博物館需求的 AI 應用，本院擬以院藏文物圖檔、詮釋資料為主要素材，結合 wiki 社群與 Taide AI 技術，製作可供博物館線上、線下應用之模型，以提高博物館近用性。

1. 廖委員鳳玓：

(1) 很肯定這項提案，透過文物研究提高 AI 生成資訊的正確性，

請問以繪畫作品作為 AI-ready data 的優先處理對象，以及相關執行過程，是否有先與院內研究人員討論？聽聽他們的想法及意見，讓他們了解 AI 特性。

2. 何委員炎泉：

(1) 過去在處理 AI 跟知識庫索引典時，皆有與數資室保持聯繫，因此大部分的研究人員都有相關的概念。我們也對於 AI 的應用有一些想法，但實際如何執行，還需等數資室有更具體的方案後再進行討論。

3. 謝委員俊科：

(1) 我們內部一直都有意見與想法的溝通，現在是試行 AI-ready data，也開始跟 wiki 這類開放社群有初步的合作，目前皆是行動計畫上的第一步。

4. 顧委員振豪：

(1) 在故宮爭取 AI 訓練上的話語權上，有關故宮的文物關鍵字與分類，應有優於 wiki 的執行方式，或是有其他自己的看法。

(2) 在 AI 資料應用方面，文字是相當重要的描述，也是目前 AI 語言訓練資料庫的重點，故宮有沒有辦法就院內所能掌控的部分，更有系統的釋出文物的詮釋資料？或是跟 Taide 做進一步需求的確認？

5. 謝委員俊科：

(1) 目前尚在摸索 AI-ready 資料的準備方式，文物的 meta data 不僅高度格式化，也經過歷屆同仁的編輯，因此這些資料應是非常接近 AI-ready，但是我們還是希望可以透過這個會議開啟行動計畫的第一步，依優先順序執行。

(2) 我們也會與 Taide 合作，並制定之後釋出資料的時程規畫。

6. 顧委員振豪：

(1) 謝謝主任的說明，我的意思類似於建置數位典藏的概念，故宮有非常多的文物，這些文物的詮釋資料、文字訊息或歷史研究紀錄，都可以針對 Taide 和 AI-ready 的部分逐步

做開放。

7. 林委員誠夏：

- (1) 剛剛故宮提的兩個案子我都非常支持。今天出席是以 CC Taiwan 主持人的身分，不管是 AI ready 還是 open data 都與 CC 授權高度相關，所以我簡單分享對這兩個主題的意見。
- (2) 有關 AI-ready 的部分，在授權方面現在故宮跟維基的合作，已經是踏上了第一步；技術方面，如果是符合相對應的格式，其實就是 AI-ready。今天故宮的低階圖像是 CC0 提供，中階圖像是 CC BY，如果故宮沒有執意各個 AI 訓練結果要做姓名標示，未來經過 Taide 訓練後，這些圖像的授權可能化整為零，不再承繼原作的表達，而只需要就訓練出處進行簡單的描述即可。如果故宮博物院可以接受這樣的處理方式，不論是低、中階的圖檔，甚至文物的文字資料，都已經算是 AI-ready。
- (3) 在 CC 領域，有加上 ND 或 NC 這些限制性的條件，就不算 AI-ready。但是只要資料進到 wiki data，原則上 wiki data 所有的資料都是用 CC0，如先建立到 wiki data 後再疏導到 wikipedia 上，那就必然是 AI-ready，而若無法採此流程處理者，則另外從權採直接投報到維基百科的模式，此時部份圖像或者不採開放授權，而被註解為合理使用。因此，在與維基合作或 AI-ready 的後續處置上，需提醒故宮要意識到哪些重要的文化資產(文物圖像的掃描或是重製)，是需高度保護的對象。今天跟維基的合作需要另設合理的使用標籤，但如果是爭取話語權，其實大部分低、中階資料與圖像，建議可用 CC0、PDM 或 CC BY 提供給維基百科，這樣維基導出的資料，才可以很自豪的說這些是 AI-ready。

8. 張委員瑜倩：

- (1) 針對第一個提案 AI-ready data，應有一些前置作業，故宮是否有相關的時程安排與短、中、長期的規劃？因為目前故宮 40 多萬筆的資料在 Open Data 與政府資料開放平臺上，是否要將這些資料一次性地匯入 AI 應用計畫？還是說有時程上的規劃，以幾個月內嘗試多少筆的方式進行，再回來討論效益？

(2) 考量到同時有好幾個 Open Data 專案在進行，以現有故宮人力以及人員能量，是否會造成負擔？

9. 謝委員俊科：

(1) AI-ready 可分為 wiki 與 Taide 兩部分。wiki 因有公民參與的部分，嘗試性較強，較像觀念上的試行，因此此部分故宮人力相對付出較多，資料數量則以少量為主。Taide 方面比較像是將大量資料轉移進去，由於都是以機器電腦運作為主，故宮傾向訂定短、中、長期目標，以「整個數位典藏為範圍，作為博物館的大腦」這個方向前進。

10. 林委員開世：

(1) 針對資料開放部分，我來自人類學博物館，經手許多原住民相關文物，對我們來說，開放非常困難，因為馬上就會牽涉到原住民特別法與當前非常敏感的族群問題。以這個角度來思考故宮的開放性，也許故宮的藏品幾乎沒有牽涉到原住民的東西，但還是多少會有一些具敏感性、宗教性的文物，也就是說有些文物是屬於密教、秘傳，是不適合開放，很好奇故宮是否有注意到相關問題，也就是說在開放之前，是否有審視過哪些文物需要非常嚴重的限制？甚至根本不適合開放？我知道故宮有辦過好幾次與原住民合作的展覽，以及故宮有許多藏傳佛教以及少數民族宗教的藏品，不一定全部開放才是對的，需要再過濾、檢視相關問題。

11. 吳委員曉筠：

(1) 就院內器物來說，我們有典藏藏傳佛教的文物，性質上屬於展示性的類型，有七珍八寶與佛塔類，另屬於祭祀類的祭器，在早期的展覽也已展示過。所以從宗教的考量上，目前的文物皆屬可開放範疇，在民眾端的數位典藏系統上基本上也都是開放狀態。

12. 王委員耀鋒：

(1) 目前在故宮的 open data 專區跟文字搜尋的規範中，有特別提到授權不包含專利權與商標權，也希望使用者在使用 open data 時，應避免讓他人誤會是本院產製的製成品。未

來關於 AI 應用的授權範圍，應依循過往 open data 的規則；AI 產生的成果，也應避免讓外界誤解為是本院產製的作品。

- (2) 行銷業務處辦理業務與院內商標授權與圖像授權的業務相關，因此我們是否有可能在開放的同時，亦增加本院商業利益介接的可行性。

13. 林委員誠夏：

- (1) 有關「原住民傳統創作智慧創作保護條例」屬於特殊法律保護範疇。故宮的素材或是原民的文化傳統素材來進行 AI 的訓練與開放資料的輸出跟提供，其實在法律狀態上並不會貶損原民傳智的特殊保護地位。以故宮而言，其多收藏具有百年歷史的文物，故而在 AI 或資料開放上的入手門檻會相對低；但在原住民博物館方面，應可與館所交流溝通，是否可釋出部分較不敏感的藏品，進行開放式的利用，以避免族群資料邊緣化。

二、議題討論：擬提請新增「文化教育」為「高應用價值主題」，加強博物館開放資料的多元化，並促進文化教育領域的資料運用。

說明：目前政府資料開放平台建置的「高應用價值主題」專區涵蓋多個領域，但缺乏文化教育面向。博物館作為文化教育核心機構，擁有豐富的資料資源，如館藏數據、展覽介紹和教育資源等。故擬提請將「文化教育」納入高應用價值資料主題，能促進相關資料之創新與教育應用。

1. 林委員誠夏：

- (1) 數位部高應用價值資料的部分，由於初期的規畫以民生經濟需求與交通、災害、水文、天候等立即性需求為主要標的，因此尚未納入文化教育領域的開放資料。此開放資料進程應是參考歐盟的演變，歐盟在開放資料施行之初，有鑒於教育文化領域本來就有行之有年的慣例與體系，因此並不直接干預文化教育領域。到 2019 年，文化教育領域（GLAM）方始列入歐盟開放資料法律管制範圍，初期仍採獎勵機制，鼓勵文化典藏機構將典藏資料以開放形式提

供給公眾，並不強迫提供。國發會應是學習到這個機制，因此初始階段列出的八大高應用價值資料，並不直接列入文化資料。

- (2) 若故宮提案要將文化教育列為第九類，基本上數發部應不會拒絕。只是若此事僅故宮執行，力量稍微薄弱，因此在簡報中也看到將召集文化部與教育部共同會商。過去曾在文化部擔任幾屆的開放資料諮詢會議委員，文化部配合執行與成本難度應沒有太大的問題。教育部方面，想請益故宮是否有跟教育部進行事前討論溝通？教育部歷來有統一教育標準的任務，因此在 CC 授權上通常皆有 ND（禁止改作）的限制。這也導致一個問題，即若有 ND 限制，就無法進入高應用價值資料專區，因為高應用價值資料專區採開放定義（Open Definition）的立場，提供民眾無限制目的、信賴性的使用。過去也協助國科會與教育部溝通，使教育部同意全字庫以開放字型授權條款（Open Font License）的方式釋出，耗時將近兩年的時間。因此可以預見教育部在這方面需要花一些時間進行溝通，因此比較好奇是否有事先與教育部溝通？教育部是否支持？

2. 謝委員俊科：

- (1) 故宮方面尚未詢問文化部或教育部，目前僅是發想階段。若是以「文化藝術」提案是否可行？因為目前看來若僅是文化部與故宮，較容易施行，教育類則等後續擴展。

3. 林委員誠夏：

- (1) 若是朝「文化藝術」提案亦可，但教育部其實也慢慢在調整，只是說其既有政策需要比較縝密的內部研議。建議先初步與教育部意見交流，確認整體推行難易度，再來調整政策方向。

捌、會議決議：

- 一、 本次所提之兩個提案，委員們皆表示支持。
- 二、 在 AI-ready data 方面，將持續推動本院資料開放之業務，目前優先以繪畫類文物為主要開放對象，並確立資料釋出的時程規畫，有系統的開放本院之文物詮釋資料；將與 Taide 與 Wiki 社群進一步合作，確立雙方對於資料開放之需求，以及後續應用的方式。

三、在增設高應用價值主題方面，將與文化部與教育部進行相關主題規畫的研議，並於後續的政府資料開放工作小組會議進行提案。

捌、散會：下午 11 時 40 分。