

...



(/)

EN (/En/)
(/Home/Sitem

博物新
知

博物館家族
(/MusData)

活動資
訊

公告事
項

線上資
源

...

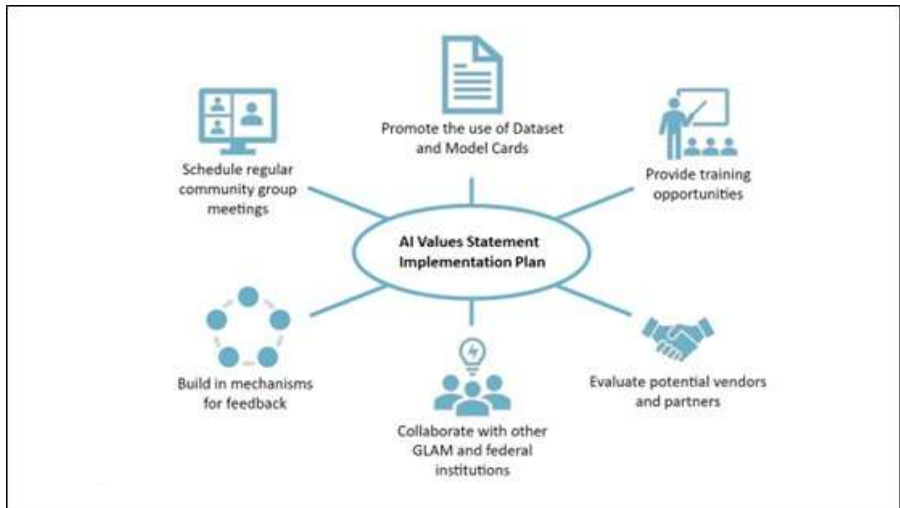
首頁 (/) / 博物新知 / 博物館科技 (/InternationalNew?Type=3)
/ 「博物館推動AI人工智慧」美國史密森學會《人工智慧價值聲明書》創制與試行

「博物館推動AI人工智慧」美國史密森學會《人工智慧價值聲明書》創制與試行

2025/04/10 閱讀數：193

關鍵字

你會喜歡的文章



「博物館推動AI人工智慧」美國史密森學會《人工智慧價值聲明書》創制與試行

(/Upload/images/77C4EF1B-8985-4DFC-8862-9151590408DB.jpg)

林聖硯（英國萊斯特大學博物館學博士）

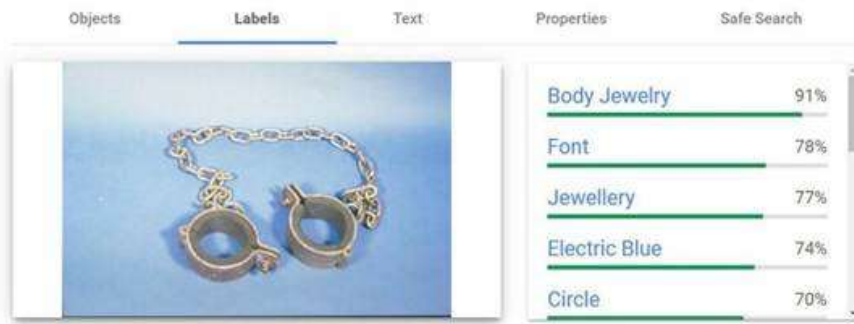
世界各地許多博物館與文化遺產組織，正在評估如何使用人工智慧(Artificial Intelligence，AI)。當前AI在美術館、圖書館、檔案館和博物館(Galleries, Libraries, Archives and Museums，GLAM institutions)的GLAM機構主要運用在「文化遺產和館藏品的管理、使用和研究」、「提升遊客體驗」、「相關業務運作與管理」。無論運用現成AI工具(off-the-shelf AI tools)以增進工作與任務執行效率；或是採用開源AI系統(open-resource AI systems)將工作自動化等

(Heritage Fund, 2023)。其中，AI在GLAM機構中最有前景的功能之一是提高數位化資料的可及性，確保所有使用者都能觸及、取用數位化內容和資料，無論是將圖像添加替代文字(alt-text)，用以描述圖像或外觀；或是將文字轉錄為音檔和影像，以及翻譯成多國語言等(Dikow et al., 2023)，透過圖像分割、物件識別、自然語言識別處理，使得數位館藏能易於使公眾搜尋、標記、使用。

個案介紹

美國史密森學會(The Smithsonian Institution)包括21座博物館、8座研究中心、15個檔案庫、21座圖書館和1座動物園，到2023年，館藏已包含約1.573億件物品和標本、15.25萬立方英尺的檔案、230萬件圖書。而在數位化藏品方面，已有3700萬件物品和標本數位化(Smithsonian, 2023a)。史密森開放取用協議於2019年2月發布，截至2024年11月已有超過490萬件線上2D和3D數位檔案採用創用CC授權(Creative Commons license)的CC0授權，釋出數位檔案給全球公眾使用(Smithsonian, 2024)。

由於COVID-19疫情，史密森學會的GLAM機構在2020年3月至2020年7月，以及2020年11月至2021年1月曾經兩度對公眾關閉，當時虛擬參訪和線上館藏查閱的需求大幅提高，進一步促使史密森學會加速將資料轉移至網路上，使觀眾能線上取用(Dikow et al., 2023)。隨著數位化步伐加快，各種預先訓練的機器學習任務也增加了可用性(availability)和易用性(usability)，儘管史密森學會希望能夠在其館藏中測試這些機器學習的模型，像是針對圖像資料庫自動分類，然而從相關研究對於圖像分類模型的基準測試結果，已被證明包含偏見、種族主義和攻擊性標籤 (Crawford and Paglen 2021, Denton et al. 2021)，且如ImageNet 資料集、Google Vision等用於視覺目標辨識的視覺資料庫，也不完全適用GLAM機構所持有的收藏品。如Dikow et al. (2023)研究中所舉例，萊瓦爾·伯頓(LeVar Burton)在故事設定為美國奴隸制時期的影集《根源》(Roots)中所配戴的鐐銬道具，被Google Vision分析處理後，可能會被歸類為珠寶。公開此藏品標籤，可能會對觀眾或特定社群造成傷害。



圖說：萊瓦爾伯頓在影集《根源》所配戴的鐐銬經Google Vision分析處理

圖片來源：Dikow, R. B., DiPietro, C., Trizna, M. G., et al. (2023). Developing responsible AI practices at the Smithsonian Institution. Research Ideas and Outcomes, 9: e113334.

<https://doi.org/10.3897/rio.9.e113334>

本圖採用 CC0 公共領域授權

有鑑於此，史密森學會於2021年成立以AI和機器學習為重點的閱讀小組，在跨組交流後，進而決定創建《人工智慧價值聲明書》(AI Values Statement)作為AI最佳實踐參考文件，以及根據《人工智慧價值聲明書》內容建議，選擇一些史密森學會資料集進行測試，並採用聚焦於自然語言處理(natural language processing, NLP)和機器學習的Hugging Face平台提供的工具模板，設計用以解釋塑造資料過程和基本原理的「資料集卡片」(Dataset Cards)，從而探索AI為史密森學會產生的影響(Dikow et al., 2023)。

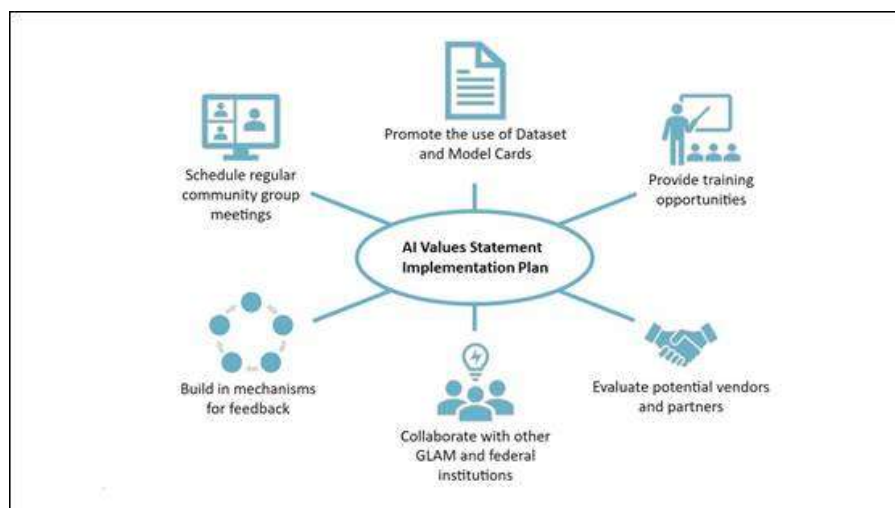
個案分析

史密森學會內部對於使用AI工具的探討分為「研究」和「策略」兩個層面。研究部分檢視研究人員針對各類資料集建立客製化機器學習模型或自然語言處理的管道，以及如何運用史密森學會館藏內容訓練和調整模型，如在數位化自然史資料集，即統整出8項運用AI的相關研究項目，包括：一、用於識別數位化植物標本紙上植物像素的分割模型；二、用於識別數位化植物標本上的汞染色，並區分型態相似的蕨類植物分類模型；三、用於識別亞馬遜魚類的屬別，以提供保護的分類基準；四、用於分類並測量數位化植物標本中李屬植物的形態變異；五、使用自然語言處理管道，以用來提取命名實體和其上下文的代名詞；六、使用自然語言處理技術，將DNA序列讀取並分類為海洋生物宿主或污染物；七、用於

識別蕨類植物及其近緣種，並評估其型態空間與物種豐富度的關係；八、成立由專門開分人工智慧的新中心AstroAI，推動下一代天體物理學(Dikow et al., 2023)。

策略層面，史密森學會評估AI對其機構產生的影響，儘管確認AI應用可在研究中提高效率，但在提升效率之外，需要關注AI應用是否有創新方式來取用數據資料和特徵，進而生成科學觀點。這類策略應用與機構的商業模型緊密關聯，因此需要更嚴謹審查，以辨別避免不準確或是不當的輸出，其團隊創建的《人工智慧價值聲明書》即是因應此目標的解決方案。2022年春季版的《人工智慧價值聲明書》從「AI為文化遺產帶來的機會」、「需要具備的考量」、「促使推動的行動」等提出建議，6項行動摘要為：參與內部社群、推廣使用資料集卡片和模型卡片、提供培訓機會、建立回饋機制、與其他GLAM和聯邦機構合作、評估供應商和合作夥伴(Smithsonian, 2023b)。

其中「推廣使用資料集卡片和模型卡片」，資料集卡片是關於資料集の後設資訊及使用方式說明，向使用者傳遞機構專業知識、背景資訊，與解釋可能有偏誤的內容。由團隊成員所完成的資料集卡片已發布在GitHub(<https://github.com/Smithsonian/dataset-cards>)上。資料集卡片上詳細說明資料集收集的初衷、背景、假設、資料變更、已發生的轉換，以及已知偏誤的解釋及其社會影響等(Dikow et al., 2023)，可協助使用者更加清楚機構特定的後設資料和編目的做法，使得機構提供的資料集更易於內部成員和公眾使用，以減少使用錯誤的風險。



圖說：史密森學會《人工價值價值聲明書》倡議的行動摘要

圖片來源：Dikow, R. B., DiPietro, C., Trizna, M. G., et al. (2023). Developing responsible AI practices at the Smithsonian Institution. Research Ideas and Outcomes, 9: e113334.

<https://doi.org/10.3897/rio.9.e113334>
(<https://doi.org/10.3897/rio.9.e113334>)

本圖採用 CC0 公共領域授權

結論與建議

隨著GLAM機構開放公眾取用數位化檔案資料，這些開源資料還包含與其相關聯的後設資料，有些後設資料是在數位化過程中自動生成，有些需要檔案管理員、編目員、館藏資訊專家、資料管理員等專家手動輸入，而將AI和機器學習工具引入工作流程，以提高效率，無疑是應用的極佳機會。史密森學會在推動AI應用上，成功結合研究與策略雙層面的探索，為文化機構使用AI樹立範例。在研究層面，透過建立分割、分類模型及自然語言處理管道等，實現資料處理效率提升與創新資料分析，展現AI在文化遺產應用中的多樣潛力。而在策略層面，全面審視AI潛在風險與偏誤，並通過《人工智慧價值聲明書》建立核心的行動框架。值得注意的是，資料集卡片和模型卡片作為工具，不僅強化資料的可用性，還促進內外部成員對館藏資料集的正确理解與使用。從軟性的AI職能訓練到硬實力的技術運用，全面檢視、規劃、試驗，以讓AI應用於文化機構能有最佳實踐的範例。

資訊來源

Crawford K, Paglen T (2021) Excavating AI: the politics of images in machine learning training sets. AI & SOCIETY
<https://doi.org/10.1007/s00146-021-01162-8>

Denton E, Hanna A, Amironesei R, Smart A, Nicole H (2021) On the genealogy of machine learning datasets: A critical history of ImageNet. Big Data & Society 8 (2).
<https://doi.org/10.1177/20539517211035955>

Dikow, R., DiPietro, C., Trizna, M., BredenbeckCorp, H., Bursell, M., Ekwealor, J., ... &

White, A. (2023). Developing responsible AI practices at the Smithsonian Institution. *ARPHA Preprints*, 4, e113335.

Francesca Collins (2024). Nemo presents recommendations for AI in museums. Retrieved November 17, 2024, from <https://www.museumsassociation.org/museums-journal/news/2024/04/european-sector-body-presents-recommendations-for-ai-in-museums/#>

Heritage Fund (2023). Artificial Intelligence: a digital heritage leadership briefing. Retrieved November 17, 2024, from <https://www.heritagefund.org.uk/about/insight/re-intelligence-digital-heritage-leadership-briefing>

Smithsonian (2023a). National Collections. Retrieved November 17, 2024, from <https://www.si.edu/dashboard/national-collections>

Smithsonian (2023b). "Developing responsible AI practices at the Smithsonian Institution" published in Research Ideas and Outcomes. Retrieved November 17, 2024, from <https://datascience.si.edu/ai-values-statement>

Smithsonian (2024). Smithsonian Open Access. Retrieved November 17, 2024, from <https://www.si.edu/openaccess> (<https://www.si.edu/openaccess>)



開啟網站導覽

數位典藏 (<https://collections.culture.tw>)

線上展覽 (<https://orange-sky-057730700.3.azurestaticapps.net/index.html>)

[Open Data介接文件 \(/Attachment/博物館API文件\(V3.1\).pdf\)](#)

[關於我們 \(/Home/Detail\)](#) [隱私權及安全政策宣示 \(/Home/Privacy\)](#)

[著作權聲明 \(/Home/Copyright\)](#)

地址：24219新北市新莊區中平路439號南棟13樓

客服電話：(02)8712-3000 (週一至週五上午9:00-12:00；下午1:30-6:00)

網站更新日期：2025.9.5



中華民國文化部 版權所有 © Ministry of Culture. All Rights reserved.