

# Wikidata 在藝術場域的應用方式

兩年研究計劃的階段性回顧 ( 上 ) 研究員余建勳

## 一、前言

在新樂園藝術空間近兩年的研究計劃中<sup>1</sup>，以 1990 年代以降的臺灣替代空間資料庫建置、及相對應的資料開放應用為主要研究目標。在經過相關需求、資源與技術評估後，本計劃認為由德國維基媒體協會 ( Wikimedia Deutschland ) 和維基媒體基金會 ( Wikimedia Foundation ) 共同經營的 Wikidata 及 Wikibase 平臺技術，在國內藝術場域有著巨大的發展潛力。對此，新樂園藝術空間透過一連串研究計劃與跨領域交流，逐步將相關技術及其應用引進國內。本計劃相信，藉由 Wikidata 及 Wikibase 的連結開放資料 ( Linked Open Data ) 特性，將在國內建立不同單位資料庫的串聯機制，且有利於後續與國際各大資料庫連結，在未來數位科技應用中奠定重要基礎。

這兩年來執行研究計劃的情境，源於新樂園藝術空間體認到當前臺灣藝術場域，在風格潮流、展示空間、生產機制、運作生態等面向已有大幅轉型。在以替代空間為行動者 ( actor ) 進行思索時，不難發現當前藝術場域已經對於不同型式、媒材或主題的藝術創作有著相對較大的包容空間。過往替代空間經常強調自己是不同於美術館、畫廊等的獨立展示空間，但迄今創作者能取得的展示機會已大幅增加，對於風格與媒材的限制也大幅減少。同時，臺灣藝術生產機制與運作生態逐漸朝向多元化與國際化的方向發展。對於自 1990 年代中葉生存至今的替代空間來說，自身機構 - 空間的生存前景則是相對越來越不明朗與樂觀，也面臨方向不一的轉型規劃<sup>2</sup>。即便如此，到了 2019 年，新樂園藝術空間已經思索到一個藝術空間最終都會面

---

<sup>1</sup> 這包括 2020 年的「新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查」( 國家文化藝術基金會藝文環境與發展補助 )、及 2021 年的「藝術資料的結構化應用：從檔案館、美術館、藝術機構操作挑戰談起」( 國家文化藝術基金會藝文環境與發展補助 )。

<sup>2</sup> 姚瑞中、吳達坤、呂佩怡 ( 2011 )。《搞空間：亞洲後替代空間》。臺北：田園城市。

臨到的問題：「若新樂園藝術空間必然走到結束空間生命的時刻，新樂園可以留下什麼既有時代意義的價值給予這個環境，作為一種參照或是典範？」<sup>3</sup>

在宣告藝術空間機構即將關閉時，至少會在兩個方面遇到相當棘手的難題：首先，既有的藝術空間檔案資料要如何保存，這包括要思考實體檔案的存放方式、數位檔案或資料庫的技術建置，及相對應的歸檔方法、內容授權、保存經費等問題；再者，這些藝術空間檔案資料將在未來如何被有效應用，這包括對於藝術空間本身的意義探討、相對應的歷史定位研究等。在當前國內，儘管在圖書館、博物館、美術館等體系均曾有過相關討論與對話，但對於沒有穩定經費挹注的替代空間 - 藝術空間來說，這類討論不一定能適用自身檔案的保存，更遑論其後續擴大應用。

本研究計劃認為，Wikidata 及 Wikibase 平臺技術將是國內藝術空間檔案數位典藏一種可能的發展方向。但對於國內藝術場域來說，Wikidata 及其相對應技術的引入，意味著不同領域間的磨合嘗試。在第一年研究計劃之中，我們大量邀請不同專業領域人士，希冀理解不同場域對於建構連結開放資料庫的應用想像，並以此成為我們在討論 Wikidata 這項技術的參照<sup>4</sup>。而在第二年的研究計劃之中，我們則是延續先前的技術討論基礎，繼續邀集不同資料研究與應用人士，協助我們設想未來的資料庫可能會被哪些不同應用習慣的人們使用著<sup>5</sup>。在這過程中，我們逐漸掌握到 Wikidata 的架構、實踐、優缺點，及我們之後未來應當朝向的方向。

## 二、Wikidata 概要

Wikidata 的構想最早源於 2005 年的維基媒體國際會議 (Wikimania) 的討論。當時便有維基用戶提到，相較於維基百科、維基文庫、維基語錄的等維基媒體計劃都是以人們作為主要的文章閱讀受眾，當時缺乏能夠有效讓電腦運算讀入資料內容的平臺<sup>6</sup>。這導致當維基媒

---

<sup>3</sup> 新樂園藝術空間 (2020)。新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查。臺北：新樂園藝術空間。

<sup>4</sup> 新樂園藝術空間 (2020)。新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查。臺北：新樂園藝術空間。

<sup>5</sup> 新樂園藝術空間 (2020)。藝術資料的結構化應用：從檔案館、美術館、藝術機構操作挑戰談起。

<sup>6</sup> 在更早之前，開放資料社群曾經藉由不同語言的維基百科條目內容，提出 Semantic MediaWiki 和 DBpedia 等計劃。這類計劃主要是藉由彙整維基百科條目的資訊框、表格等處的內容，另外建立可供參照的連結資料集，某種程度可以說是 Wikidata 的前身。而即便 Wikidata 成立後，部分計劃仍然持續進行著。

體計劃所收錄的內容越來越龐大的時候，反而會導致一些問題的解答反而更難直接取得，例如「某個國家最多居住人口的城市」、「某個國家特定地方層級的女性首長」等等。這就像 Wikidata 產品經理莉迪亞·平徹 ( Lydia Pintscher )<sup>7</sup>在接受我們訪問時所形容的，「當時這些資料是被『鎖』在維基百科的」。平徹提到，過往的維基媒體計劃要同時間更新相同主題的基本資料 ( 例如最新職位、逝世日期等 ) 時，往往必須要由用戶逐一至不同語言版本的相應頁面中進行修改，而無法藉由單一編輯同時完成所有平臺內容的更新<sup>8</sup>。

為了有效地讓資料能夠被電腦程式讀取、及簡化對不同語言版本內容的修正流程，串聯不同維基媒體計劃的 Wikidata 在 2012 年 12 月正式成立。Wikidata 計劃的開發與研究主要是由德國維基媒體協會負責，而為基媒體基金會主要負責該平臺的伺服器營運工作。不過隨著 Wikidata 自身所收錄的資料不斷擴大，其原先的功能便不再單純只是作為維基媒體計劃的跨語言連結平臺，而是集合不同領域、語言結構化資料的超大型連結開放資料庫<sup>9</sup>。

大體上，Wikidata 的資料都是由編輯用戶對內容進行維護。其中一部分是由個別用戶以不同技術方式提交與編輯資料，並因此參與國際 Wikidata 社群之中。根據統計，目前每天大約有 12,000 人會編輯至少 5 筆數據項。平徹提到一種區分維基媒體運動用戶的方式：一部分用戶確實更加享受撰寫維基百科條目的樂趣，並願意對於世界的某項主題進行更深入的研究，並從頭到尾撰寫出完整的主題知識內容；另一部分用戶本身對於撰寫維基百科條目的意願不大，並有對於維基媒體運動的其他想像，例如希冀藉由每一筆微小貢獻去成就更為宏觀的結果，或者是更有興趣以結構化資料的方式整理和描述世界樣貌，因而這些用戶往往會更願意投入 Wikidata 的工作。對於執行藝術主題來說，往往同時需要這兩種類型的用戶共同合作。而對於比較偏向以條目 - 文章敘述世界的用戶來說，藉由 Wikidata 提供的資料視覺化工具呈現在該條目或文章中的研究主題，往往能夠更有效地讓他們理解 Wikidata 強調結構化資料的重要性，並能促成更多結合質性研究與量化研究的合作可能性<sup>10</sup>。

在另外一部分案例中，主要是由不同組織機構願意與啟動 Wikidata 的合作計劃，並且向 Wikidata 貢獻自身所收錄的資料內容。這些組織機構有的是基於其特定專業背景或主題領

---

<sup>7</sup> 莉迪亞·平徹 ( Lydia Pintscher )，德國電腦科學家、Wikidata 產品經理，也是德國維基媒體協會成員。

<sup>8</sup> 線上訪問莉迪亞·平徹，2021 年。

<sup>9</sup> 新樂園藝術空間 ( 2020 )。《新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查》。臺北：新樂園藝術空間。

<sup>10</sup> 線上訪問莉迪亞·平徹，2021 年。

域所組成的美術館、博物館、檔案館等機構，有的便直接是涉及到許多資料紀錄的政府機構或教育部門等，而有的則是基於特定宗旨、倡議或目的而成立的非政府組織、民間單位、資料庫等。最常見的機構合作案例，便是將該機構原先已經自行建置完成的資料庫資料內容導入 Wikidata 之中。Wikidata 的團隊或社群通常會協助有意願合作的組織機構，包括找尋相關專業人士、相關計劃資源、對相關主題有興趣的編輯用戶等，並確立其在整個維基媒體運動中可以協助或貢獻的角色。平徹便曾經提到，由於許多對特定主題抱有興趣或專業的組織機構不一定能夠馬上理解、融入維基媒體運動之中，因此 Wikidata 計劃團隊和社群必須盡可能促成該組織機構與 Wikidata 之間的合作，使新加入的組織機構也能加入成為整個維基媒體運動的一員。同時在藝術領域這個部分，許多對於藝術主題有興趣、或者過往曾經參與合作計劃的 Wikidata 用戶，已經組建關於藝術主題的社群或計劃，並與許多同樣抱持著相同目的的組織機構展開合作<sup>11</sup>。

### 三、藝術場域應用方式

發展至今日，本計劃認為可以將 Wikidata 總結為一個收入多語言、多領域結構化資料的超大型資料庫，且該資料庫的內容免費地開放給所有人使用。依照 Wikidata 的著作權規定，這些結構化資料採用「創用 CC 公眾領域貢獻宣告 (CC0) 1.0 授權條款」，意味著這些資料被視為是已經進入到公有領域 (Public domain)。所有用戶或機構在向 Wikidata 提供或貢獻資料時，均要認識到已經放棄在著作權法規之中所享有的著作權利；對於使用這些資料的用戶或機構來說，這意味著他們能夠在各種不同場合下，完全不需要任何事前或事後的許可，相當自由地去複製、修改、分享、展示、發表或運算這些資料內容，也可以藉由這些資料內容進行商業目的的開發利用，而不需要提供任何著作權費用<sup>12</sup>。

本計劃在第一年研究調查中，已經詳細地介紹了 Wikidata 所收錄的資料架構。簡單來說，Wikidata 的數據項 (Items) 代表其收錄的不同資料主題，每一個數據項都代表著一個特定的主題 (或者是特定維基百科的頁面)，並會賦予一個唯一由該數據項使用的項目標識符，由前方英文字母「Q」加上一連串數字組成 (例如「保羅·塞尚」的數據項標識符為「Q3554

---

<sup>11</sup> 線上訪問莉迪亞·平徹，2021 年。

<sup>12</sup> 創用 CC。CC0 1.0 通用 (CC0 1.0) 公眾領域貢獻宣告。 *Creative Commons*。 [https://creativecommons.org/publicdomain/zero/1.0/deed.zh\\_TW](https://creativecommons.org/publicdomain/zero/1.0/deed.zh_TW)

8」)<sup>13</sup>。另外，維基數據同樣會為作為不同資料性質內容的屬性 ( Properties )，建立如數據項一般的獨立頁面，由前方英文字母「P」加上一連串數字組成 ( 例如用來說明主體就讀教育機構的「就讀學校」屬性，其屬性標識符為「P69」)。但是不同於前面的數據項允許用戶自行建立和編輯，屬性必須經由維基數據社群的批准後，才能創建該屬性、並應用在其他數據項頁面之中<sup>14</sup>。關於特定數據項的詳細特徵，主要是由頁面下半部分的聲明 ( Statements ) 群組建構其更為豐富的意義內容。這些聲明群組採用鍵-值對 ( Name-value pair ) 的資訊形式，因此每一項聲明都會提供特定的屬性，以及連結該屬性的值。

在第一年的研究調查中，我們提到 Wikidata 收錄資料的方式其實類似於維基百科，都是參考既有的文獻資料之後，對於這些知識內容提取整理而成。隨著 Wikidata 的平臺運作機制越趨完整，通常能夠在 Wikidata 建立數據項者，大致可以區分成三種情況：(一) 在維基百科等維基媒體計劃之中存在著相對應的頁面；(二) 能夠藉由公開的可靠參考文獻佐證之資料內容；(三) 基於資料庫內容結構和邏輯，需要單獨建立的特殊數據項。值得注意的是，由於 Wikidata 要求所有資料內容都應該附有可靠的參考文獻，這意味著其不適合成為藝術機構首次公開發表資料集的平臺，因此許多藝術機構、單位的展覽若未經過公開出版的過程，並無法直接收錄在 Wikidata 之中<sup>15</sup>。

平徹在訪談中提到，對於大型機構或政府單位來說，將自身單位所持有的資料貢獻、開放給更多的人，意味著更多人能夠運用這些資料<sup>16</sup>。對於藝術機構來說，如果特定單位選擇將自身擁有的資料導入 Wikidata 之中，這些資料將能夠與 Wikidata 既有的大型資料集彼此相連，能夠擁有更為豐富、多元的運用可能性。例如，在將有關臺灣藝術家的連結開放資料導入 Wikidata 中後，這些資料便會因為連結開放資料架構的獨特性，與其他來自不同地區、機構、單位提供的藝術資料相互串聯，甚至是與原先分屬於不同主題、專業的資料集合產生聯繫。另外一種可能的應用是，在同時擁有藝術家資料庫、及地理資料庫，並將兩者資料均導入 Wikidata 之後，由於許多藝術家作品、展覽等主題或名稱往往與特定地理位置有關，便有機

---

<sup>13</sup> 新樂園藝術空間 ( 2020 )。《新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查》。臺北：新樂園藝術空間。

<sup>14</sup> 維基數據 ( 2022 )。Help:Properties。Wikidata。 <https://www.wikidata.org/wiki/Help:Properties>

<sup>15</sup> 新樂園藝術空間 ( 2020 )。《新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查》。臺北：新樂園藝術空間。

<sup>16</sup> 線上訪問莉迪亞·平徹，2021 年。

會查詢 Wikidata 資料庫裡面的資料，更快速地彙整出藝術家作品、展覽與特定地理位置的關係，並有利於後續的策展、評論、研究等應用。

## 四、代結語

無庸置疑地，如果要另外建置如 Wikidata 這樣的連結資料庫，需要花費相當大量的研究與建置成本，且難以確保該平臺的後續營運與維護。雖然即便是運用如 Wikidata、Wikibase 等既有的平臺，資料內容本身的內容彙整、格式整理、上傳資料本身也會花費一定的精力與時間，但是其後續產生的效應是相當豐富的。其中，在針對相當大量的文獻資料內容進行分析運用的時候，過去一般是由單一個人或團隊逐一從大量的參考文獻之中，艱困地尋找、整理或盤點出特定的資料內容。且如果在不久之後，有需要在相同範圍的文獻資料中爬梳其他主題的資料內容時，就必須以人工方式再度翻查先前閱讀的參考文獻，重新尋找相對應的內容。

但是相反地，在將大量資料內容藉特定格式整理出相對應的結構化資料後，並上載至諸如 Wikidata、Wikibase 等平臺上，由於這些資料庫系統提供完善的導出功能，因此能夠更便於其他研究人員、第三方人士應用。由於這些資料內容能夠透過電腦進行讀取判斷，研究人員便可以利用不同資訊檢索技術分析這些資料內容，甚至依照不同工作目的、研究主題打造出專屬或通用的資料視覺化工具。這意味著其他研究人員或第三方人士不需要重複查詢、驗證資料正確性的工作，可以將重心放在對於這批結構化資料的處理，並以各種有創意、或更為深刻的方是運用這些資料。而當這些不同主題的結構化資料依照特定方式進行串聯時，將更方便研究人員、使用受眾找出過往人工查詢方法不易發現的問題意識與研究內容。

在與 Wikidata 產品經理平撒的訪談中，我們注意到推廣 Wikidata 的過程之中，其實也面臨到來自不同背景人士對於該計劃的各式想像與意見。例如在 Wikidata 計劃之中，必須同時面對維基媒體運動社群、Wikidata 技術開發團隊、及內容生產者和受眾<sup>17</sup>。在某種程度上，這意味著 Wikidata、Wikibase 平臺技術、甚至是相關的連結開放資料計劃的推動，都必須試著有效地結合內容生產者 - 受眾和技術團隊兩方的專業，並共同解決需求與問題。在面對美術館 / 博物館 / 藝文機構等單位提出捐贈合作連結開放資料時，Wikidata 或 Wikibase 資料庫團隊的串聯與溝通能力便顯得相當重要，必須確保合作雙方能夠理解相關工作的執行方式、開放授權、及其往後應用等資訊。

---

<sup>17</sup> 線上訪問莉迪亞·平撒，2021 年。

但是，由於 Wikidata 要求收錄的結構化資料需要附上參考文獻，這導致其不適合作為資料集首次公開的地方，而必須先另行在線上公開或實體出版<sup>18</sup>。對此，我們注意到假若運用作為 Wikidata 基礎架構的 Wikibase 平臺技術，藝文單位便能夠獨自建立類似 Wikidata 的客製化連結開放資料庫，並能自行決定開放的內容。因此，在「Wikibase 在藝術場域的未來展望：兩年研究計劃的階段性回顧（下）」中，我們將特別剖析藝術資料的獨特性、及 Wikibase 可能可以處理的面向。

---

<sup>18</sup> 新樂園藝術空間（2020）。*新樂園 25 年藝術資料的結構化想像與藝術空間檔案研究調查*。臺北：新樂園藝術空間。