

擴增研究：臺灣人物傳記資料庫 (Taiwan Biographical DataBase) 工作坊小記

文——林佩蓉 研究典藏組

攝影——丁千惠、吳禹中

從本期開始，我們希望陸續為大家介紹臺文館正在進行的數位人文研究工作。自1990年代起，當社會資產已進入數位化的時代之際，科技與人文的結合研究隨之發展，近年來有越來越多以科技整合為核心的數位人文研究及課程，人文研究視角與方法的擴充及改變，是數位時代下的成果。



「你被Google鎖定了嗎？」(圖片來源：http://bestlink.in.ua/novyny/files/062015/79__1.gif)

「你被Google鎖定了嗎？」

常言道：Google用得好，報告沒煩惱。這幾年來Google搜尋引擎宛如站在眾星拱月的舞臺上，受眾人的膜拜，名曰Google大神，他隨意走進人們的螢幕裡、處處受到歡迎。

Google留住了每一個膜拜者的足跡，成為難以計量的大數據，他可以存而不用，也可以隨時為任何一種調查服務，存取的方式、累積的空間、運算的方式等，芸芸眾生者，難以親近，更遑論了解箇中的原理或關鍵。對於只要答案的人，搜尋的機能、結構是什麼並不那麼重要，但對於想要擴增答案之外的成果、想要跳脫Google餵養的訊息以及跳脫他的挾制、想要真心了解自己所求所問的內容如何形成……，Google就不能被當作大神，而是要在他所運用的技術與平臺上，學習建立可自行掌握、運用的系統。

從土法煉鋼到彈指之間，Google退散

本文一開始去追問個人使用Google去求得知識的用意，是為了從本期開始，我們希望陸續為大家介紹臺文館正在進行的數位人文研究工作。自1990年代起，當社會資產已進入數位化的時代之際，科技與人文的結合研究隨之發展，近年來有越來越多以科技整合為核心的數位人文研究及課程，人文研究視角與方法的擴充及改變，是數位時代下的成果。2017年5月12日，國立臺灣文學館邀請同在進行數位人文研究的學者專家，包括清華大學祝平次教授、臺灣師範大學張素玢教授、中央研究院臺灣史研究所王麗蕉主任、張隆志副研究員、顧雅文助理研究員，以及臺文館與成功大學、臺中科技大學共同合作的「數位全台詩整合型計畫」主持人施懿琳教授、許明珠助理教授等人，共同召開「臺灣人物傳記資料庫TBDB工作坊」，以臺灣文學與歷史人物的數位

研究技術、資料庫建置經驗及成果為主題進行交流，同時邀請青壯輩學者，分享資料庫與人文研究的關聯，從土法煉鋼到彈指之間，在這樣的討論中，共同的核心在於「人」，包括對人的研究及所產生的數據（data），「人」設定在作家及史學家、知識分子類別裡，有其時間的脈絡、亦有其人際關係的溫度。在這場工作坊中，透過蔡明諺老師、陳淑容老師、柯榮三老師、向麗頻老師分享多年來查考資料的方法論，我們了解在多種文獻中互文比對，為問題意識尋找線索與解答，運用史料想方設法，是研究者的功力，即使是土法煉綱，也是重要的基礎功，是走向彈指之前，必備的技能。

我們能夠想像，這些學者在研究室裡，爬梳田調所得的史料、建立匯整表格檔案、查找資料庫進行比對，他們正在描繪以「人」為核心的知識系譜，因為發現問題，然後著手尋求解答，Google退散。

擴增研究，TBDB的進行式

祝平次老師以「CBDB（中國歷代人物傳記資料庫）與TBDB（臺灣人物傳記資料庫）」為題，強調資料庫應作為當前研究方法論，關於CBDB資料庫的形成，在Google中有簡潔扼要的說明，本文不再敘述，但TBDB對於臺灣文學領域而言，尚屬新興之學，在概念上，同樣在是「群體傳記學」的研究，可同時對臺灣歷史、文化、文學提供寬廣的視角，如何將這些視角所發出去的「線」，聚集在同一個平臺上，運用數位技術，進行觀

1. 蔡明諺老師分析《台灣新生報·橋副刊》主編歌雷的生平經歷及與該副刊的關聯性。
2. 陳淑容老師日治時期文學文獻探勘經驗，探討知識分子的思想活動，所運用的材料包括《台灣新民報》等日治時期報紙。
3. 柯榮三老師就「清季職官表查尋系統」等資料庫，進行臺灣古典詩人的作品及行跡研究。
4. 向麗頻老師以詩社為研究對象，探討詩人透過結社，在社會網絡中的活動。



察、運算、統計，讓數據說話，是TBDB要做的事。

捷克籍的臺灣文學研究者路丹妮老師(Táňa Dluhošová)在「從數位人文談臺灣文學研究新領域——以『斷裂』及『連接』為案例」中提問：文學史的時期如何界定？其「斷裂」與「連接」的引證是什麼？什麼是「重要性」？繪製戰後初期文學場域樣貌的可能性？以及如何以社會網絡分析研究臺灣文學等。這些事實上都是大哉問，但一位對臺灣文學超級熱誠的外國學者，透過數位技術的使用，受CBDB的啟發，企圖從人物的集體研究中設定核心問題，包括「出生、死亡、婚姻、家庭、社會出身、經濟地位、居住地、教育、個人財富與來源、職業、宗教、宦歷等方面」，再將此研究範圍中所有人物的各類信息加以羅列、組合，通過對這些信息的考察找出具有顯著意義的變量。研究者可以檢測這些信息的內在相關性，及其與其他行為形式或行動形式的相關性。而核心問題，會隨著目標群體的不同，而有所設定，這也就是TBDB必須建置的原因之一。

臺灣師範大學臺灣史研究者張素玢老師，擁有很多年的地方人物史的田調經驗，也參與方志的編修，此次她以《新修彰化縣志·人物志》為底本，創造「突破人腦的極限」之可能，並尋求資訊工程與圖資學者的合作，是跨領域的結合。會中張老師也繪製了「TBDB建置流程」，以文本分析及數位分析相互合作，分層建置，包括「資料層」、「資料處理層」、「服務層」，層層都充滿挑戰。

標準化、數位化、文學人與IT人

一個友善、體貼使用者的資料庫，一定經過相當嚴謹的資料標準化過程，中央研究院臺灣史研究所的檔案館，在創所所長許雪姬教授帶領下，結合圖資專業的王麗蕉主任，這幾年完成了許多重要的「人」為主的資料庫，可見於「臺灣史檔案資源系統」、「臺灣日記知識庫」、「臺灣總督府職員錄系統」等，在顧雅文老師的介紹中，我們

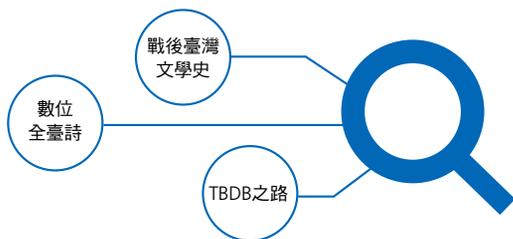


1. 祝平次老師演講，讓電腦程式、數位技術，大顯「關聯」的多樣性。
2. 路丹妮老師從「臺灣戰後初期的文學場域生成」命題出發，她以1949-50年代初期的文學雜誌為材料，逐筆輸入作家作品，逐一考證作家身分，而後透過程式運算，繪出照片中的圖片，這是其中一種數據樣態的呈現。
3. 張素玢老師以「建置TBDB的概念及構想：以《新修彰化縣志·人物志》為起點，以突破人腦極限以及不可能的任務，說明TBDB建置的要素。

看見資料庫的親民性——這些資料庫也是現今社會新聞的記者愛用的網路資源。參與工作坊綜合座談的張隆志老師，則以大眾史學的角度，強調人文領域資料庫的合作，有助於TBDB的建置。

本次工作坊另一個功能，即是將兩個原本看似不相關的環扣巧妙地連結。圖資專業、文訊雜誌社資深編輯吳穎萍老師在文訊參與文學工具書編纂的歷程中，說明了資料建立與整理方式；臺南大學數位學習科技系的林信志老師則以「資料庫知識探索」為目標，提出「資料整合、轉換與化約／選擇資料探勘的分析目的與問題類型／執行資料探勘的分析工具／結果視覺化，解釋與描述／運用探索出來的知識」，並提出「雲端平臺、知識探索與大數據分析」的技術及研究方法，我們在接下來的幾期中，將再介紹這些Google也在使用的科技運算，當然，我們會以文學為圓心，分享跨界合作所散發的能量。

打開視界：一張圖、一場工作坊



圖片引用來自工作坊計畫中之報告內容及<https://unwire.pro/wp-content/uploads/2016/01/Google-Search-590x214.jpg>

一場工作坊，15位學者與會，集結了臺灣文學、臺灣史、中國文學、圖書資訊、檔案管理、數位科技等學門，這樣的組合，跨界合作、跨領域研究，已是進行式，然而群組長期的合作，以TBDB建置為例，仍然有許多困難橫在眼前，還有許多環結需要打通，在這之前，我們希望集結更多正在進行人文資料庫的各方好友，一起打開視界，擴增臺灣文學研究。✎



1. 顧雅文老師說明中研院臺史所資料庫的關聯使用方式。

2. 吳穎萍老師以圖資專業投入文學資料之蒐集、編纂多年，介紹當前文學工具書資料庫的內容。

3. 林信志老師是協助臺文館整合十多種資料庫，建立友善使用的關鍵人物。他也帶領學生，跨越資工與人文的藩籬，結合兩方的優點，創造共享共創的平臺。