

數位人文趨勢與 「21 世紀的史家 技藝」工作坊

陳詩沛 美國哈佛大學費正清中國研究中心博士後研究員

涂豐恩 美國哈佛大學歷史與東亞語文學程博士生

何浩洋 英國倫敦國王學院歷史系博士後副研究員

近年來數位化史料和各類線上工具不斷湧現，然而在臺灣，對於歷史學與廣義的人文學界，有關數位化所帶來的問題與機會，卻鮮少公開地討論。2012 年 11 月 28 日國立臺灣大學數位人文研究中心，舉辦「21 世紀的史家技藝：數位時代的資料、工具與想像力」工作坊（The Historian's Craft in the 21st Century (Taiwan) – an informal conversation about technology and research methods for humanists，以下簡稱 THC21-Taiwan），正是希望創造一個機會，讓有興趣的研究者能集思廣益，分享彼此的問題與理念。活動分成上下午兩大部分，上午為公開論壇，下午則有四場教學講座，涵蓋「全文資料標記」、「全文資訊自動擷取」、「視覺化工具」與「地圖與 GIS」等四個主題。

這一兩年在臺灣似乎有越來越多人注意到「數位人文」，（註1）但與 THC21-Taiwan 類似的活動似乎仍算罕見。作為第四屆數位典藏與數位人文國際研討會的會前會，本工作坊有別於大型研討會的形式，著重面對面的討論，以及實際操作導向的技術課程。工作坊當天，超乎預期地有近 40 位與會者，包括資深學者與年輕的研究生，會中討論熱烈。工作坊的投影片與錄影，目前皆已公布於會議網站（<http://thc21tw.blogspot.com/>），有興趣的讀者可前往參考。以下我們想從規劃者的角度，記錄本次活動的緣由、工作坊當天狀況，以及會後的反思。

如同我們在會議網站上所提到的，任何在當代從事歷史研究的人，很難不注意到迅速膨脹的數位資料。「在這個時代，做歷史研究似乎前所未有的容易，又是前所未有的困難——有人稱之為富裕的兩難。樂觀的學者歌頌未來無限的機會，悲觀的學者洞察箇中的陷阱；皓首窮經的老師抱怨學生不再讀書，困惑迷惘的學生則在一個又一個的資料庫中暗自摸索。」（註2）

但時代的演進似乎不會因著我們的樂觀或悲觀而改變，更重要的或許是思考如何應變。過去我們曾分別在臺大資工系項潔教授所主持的「數位人文研究中心」（原「數位典藏研究中心」）與「數位典藏與自動推論實驗室」工作，對於數位技術在歷史研究的應用，共同做過實驗與嘗試。2011年起，因緣際會下，我們又一同前往哈佛大學進修，並於該地結識前來訪學的清華大學中文系祝平次教授。祝教授對於數位人文素有興趣，也有許多實作經驗。我們都感受到數位人文浪潮即將來襲，也認為我們過去在臺灣的嘗試，和在國外的見聞，或許值得分享給更多人，因此有了本次的工作坊。

除此之外，還有幾個重要的活動特別啟發我們，包括2012年1月份在芝加哥召開的美國歷史學會年會、同年1月起在哈佛校內發起的The Historian's Craft in the 21st Century雙週討論會、以及2012春季的Digital Humanities 2.0課程。以下便簡短介紹這三個活動的內容。

1. 第126屆美國歷史學會年會（American History Association Annual Meeting，以下簡稱AHA）

美國歷史學會年會是一年一度、全國歷史學者聚集交流的盛會。2012會中，有兩百餘場的論文發表、演講、工作坊、學術徵才等活動。這年的AHA特別規劃了23場與數位技術相關的場次，邀集美國歷史學者共同探討數位方法在歷史研究及教學上的應用，並以The Future is Here: Digital Methods in Research and Teaching in History為名，串連起這些場次共同的主題[<http://www.historians.org/annual/2012/digitalhistory.cfm>]。當中有幾個場次特別受我們注意，包括Digging into Data、New Directions in Spatial History、與Pioneers Discuss the Future of Digital Humanities。

New Directions in Spatial History場次邀請美國重要的幾個專長於歷史地理的團隊前來報告最新成果，其中史丹佛大學的The Spatial History Project[<http://www.stanford.edu/group/spatialhistory>]發展了一個能結合時間空間、呈現歷史資料的工具（圖1），哈佛大學的Center for Geographic Analysis介紹了他們開發的WorldMap平台，能讓研究者社群在此平台上分享圖層，套疊他人分享的圖層產生新的地圖（圖2）。

Pioneers Discuss the Future of Digital Humanities場次則邀請了Google N-Gram Viewer的開發團隊Culturomics，介紹

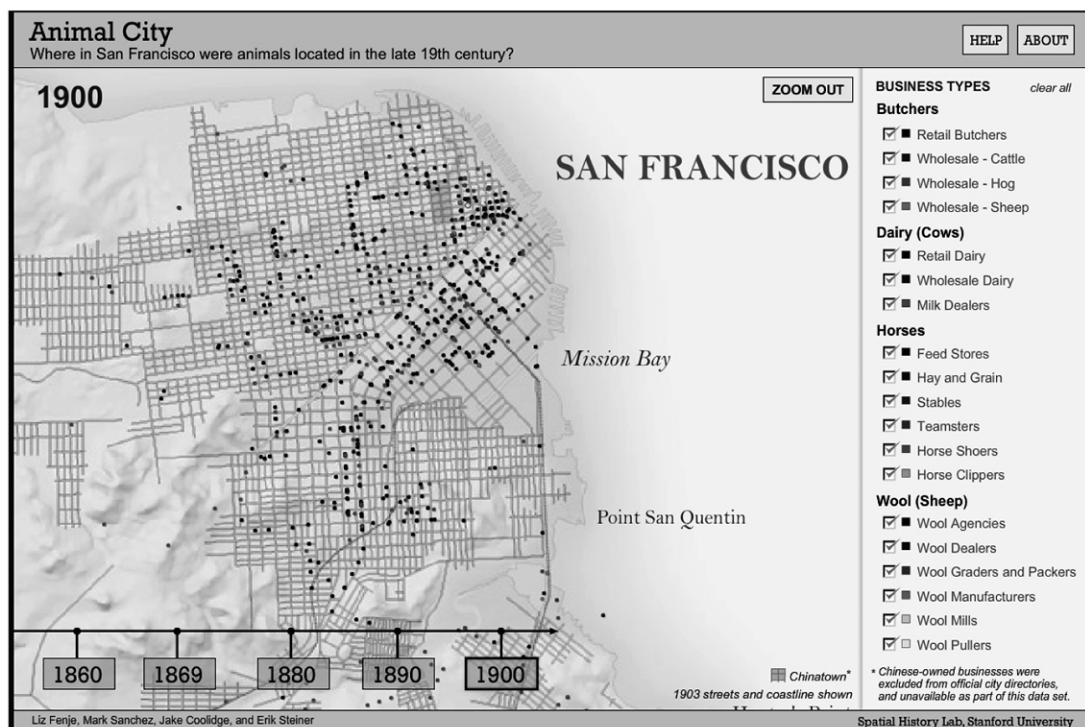


圖 1 1860-1900 年舊金山市內動物相關產業（屠宰業、乳品業、養馬相關與羊毛織品產業）的位置分佈，以時間序列呈現

他們為 Google Books 建置的最新工具 -- Bookworm [<http://bookworm.culturomics.org/>], 使用者除了能下關鍵詞觀察詞彙的年代趨勢之外，現在還將年代趨勢限定在特定領域的書籍，進行更細緻的觀察（圖 3）。

Digging into Data 是美國 National Endowment for the Humanities (NEH) 為了促進數位人文方法論的發展而設立的獎項，從 2009 年起每兩年徵選一次、每次選擇八

到十個計畫。迄今已有英國、加拿大、荷蘭加入集資，共同補助各參與國的數位人文計畫。該場次邀請其中幾個獲獎計畫前來報告成果。Old Bailey Online [<http://www.oldbaileyonline.org/>] 將資訊領域的 data mining 技術應用在歷時共 239 年的英國罪犯法庭開庭紀錄上，Mapping the Republic of Letters 則將文藝復興時期歐洲重要思想家的信件往來網絡呈現在地圖上（圖 4），Railroads



圖 2 WorldMap 中一份公開分享的中國歷史地圖，圖中套疊的圖層有：1820 年縣界（黃線），及明清時期的驛站（紅點）與驛道（紅線）

and the Making of Modern America 嘗試結合空間與時間兩種軸度，來對歷史資料進行分析與呈現，而 Towards Dynamic Variorum Editions 則試圖透過資訊方法，找出經典文

獻在各種語言中被重述、引用的軌跡。這些計畫都是針對大量人文資料，提出如何結合資訊技術、拓展新研究方法的嘗試。

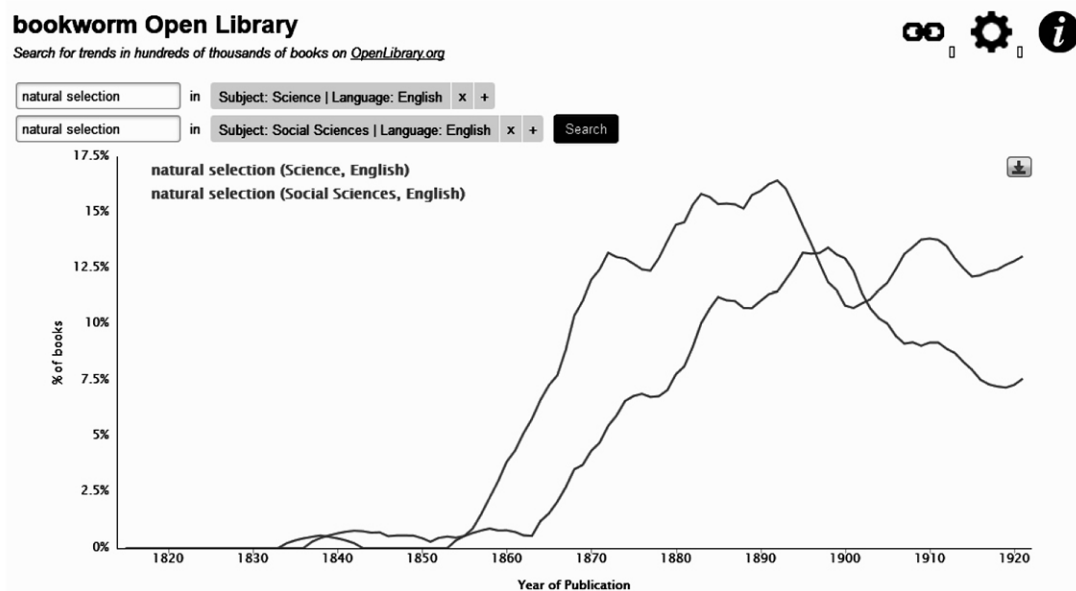


圖 3 於 Open Library 的科學類及社會學圖書中分別查詢 natural selection 一詞，以觀察達爾文的「物競天擇」理論何時開始被社會學者採用，應用於社會學理論

2. 哈佛大學 The Historian's Craft in the 21st Century (THC21) 討論會

這個系列討論會是由 Center for History and Economics 博士後研究員 Rachel Leow、歷史系博士生 Konrad Lawson、以及東亞系博士生 Javier Cha 共同發起的，正如活動名稱所言，討論會關注的是數位時代下史家應具備的技藝，在傳統的定義外，數位工具的應用儼然成為 21 世紀史家也應具備的新技藝。這個討論會是由發起人先規劃了一系列主題，邀集有興趣的人文學者每兩週見面

一次，針對：數位閱讀、數位寫作、蒐集研究材料與筆記、數位教學、使用資料庫與大量材料、視覺化呈現等主題，分享好用的工具、或者介紹自己在眾多工具間摸索的痛苦經驗。這個系列討論會獲得校內博士生與博士後研究員的熱烈迴響，也不乏有教授參與。他們並邀請與會者以 Google Drive 文件，「多人共寫」來記錄討論的內容，更是有效利用數位工具的一種實現。THC21 的討論內容筆記都已上網，可供查閱 (<http://idlethink.wordpress.com/thc21/>)。



圖 4 文藝復興時期思想家的信件往來網絡，左上圖用地圖呈現信件的收發地點，左下的年代分佈圖能進一步將信件範圍限定在特定的年代區間，右表則可以進一步限定通信者，並看到所有通信地點的統計表

3. 哈佛大學 Digital Humanities 2.0 課程

這些課程是由 Jeffrey Schnapp 教授與訪問學者 Dennis Tenen 共同主持，Schnapp 同時也是哈佛校內 MetaLab 的主任。何以當時還未普及的 digital humanities，會被逕自「升級」成了版本 2.0？Schnapp 解釋道：身為中世紀書籍史的學者，他們關心的不僅是「數位工具能為人文研究帶來什麼改變」，更是「數位技術如何改變了以往知識傳播的方式」。的確，今日知識傳播的方式

已與二十年前相異甚大，更多的人際溝通與思想交流，是透過電子郵件、網站、部落格、臉書（facebook）、推特（twitter）等媒介達成。一個明顯的例子，是人文學者在撰寫專書與論文之餘，也開始經營部落格，記錄社會觀察與研究理念。課堂中兩位老師指定了大量的文獻閱讀，帶領大家共同檢視過去幾十年來（甚至在數位人文一詞出現前），文學、歷史學、社會學等領域的學者，就已思考過的相關議題。詳細課程大綱

與指定閱讀，請見：https://docs.google.com/document/d/17wouGX_4WC0eQlNoEqvvSkQoxj607ts2tQyIMrnkUP4/edit。

以上只是我們眼見所及、美國數位人文發展的冰山一角。除此之外，至少還有好幾個數位人文的重要據點，未能在此深入介紹，譬如 University of Nebraska Lincoln 的歷史系、University of Virginia 的 IATH（The Institute for Advanced Technology in the Humanities）、George Mason University 的 CHNM（Roy Rosenweig Center for History and New Media）、馬里蘭大學的 MITH（Maryland Institute for Technology in the Humanities）等等，而加拿大、英國、歐洲等地也都有著不同方向的發展，尤其是早期數位人文發展重點的 Text Encoding Initiatives（TEI）文本標記語言，早已在歐美相當普及。（註3）

值得注意的是，上述活動都是探討數位技術與人文研究之間的關係，卻各有焦點，正反映「數位人文」一詞所引發不同的投射與反思。同時，幾乎就在這一兩年，數位人文在美國快速發展起來。一個原本鮮為人知的名詞，突然之間成為人人口中都在談論的議題。相關的工作坊，開始在各地舉辦，（註4）不少大學也成立數位人文研究中心。最重要的是，這兩年的職缺公告中，許多單位都希望能尋找有相關經驗的教員，甚至指名要「數位人文」的學者，彷彿宣告一個新的領域已經誕生。

也是在這樣的背景下，我們開始思考如何規劃以數位人文為主題的工作坊。如前所述，過去幾年臺大數位人文研究中心已定期舉辦相同主題的國際研討會，同時也有各種資料庫相關的教學活動。不少大學圖書館也會舉辦類似的資料庫研習會。但在極為宏觀的研討會，與極為特定的教學課程之間，我們隱約感覺還有一些事情值得做。比如，一種介於其中，可以同時結合理論層面與操作經驗的工作坊。

THC21 討論會的發起人 Konrad Lawson 曾提出這樣一個想法：數位人文可以分為大寫的 Digital Humanities 與小寫的 digital humanities。前者指的是立基於充足人員與經費，大規模的研究活動，這樣的活動，在各種學術贊助單位的支持之下，已越來越多。但顯然，並非每個研究者都有機會能主持或參與其中。事實上，也不需要每個人都投入這樣的研究，因此有了小寫的數位人文。小寫的數位人文，是以個別研究者為中心，將簡單卻極有效益的數位技術，應用在他們自身的課題中。換言之，大寫的數位人文是從高處出發，小寫的數位人文，卻是由下而上，希望將技術普及化。後者正是 THC21-Taiwan 的目標。

當天上午場次為公開的座談會，以三場報告為開場。陳詩沛以她參與的中國歷代人物傳記資料庫（CBDB）為例，介紹美國的數位人文發展趨勢。清華大學祝平次教授強調人文學者的主體性。他認為，人文研究與

數位技術之間應該是互相補充，而非相互競爭。素有訓練的學者，對於資料的解釋，是永遠無法被取代的。他自己在群體傳記學與社會網路分析等課題的經驗，正是絕佳的實例。涂豐恩將數位人文的發展，置於全球視野和公共視野兩個脈絡之下。他認為，數位技術既可以讓人文研究跨越許多限制，同時也可以打破學院內外的區別。研究者因此應該思考如何可能與非專家對話，將研究成果呈現給大眾。

在場來賓對於這些主題，也予以熱烈回應。作為東道主的臺灣大學數位人文研究中心主任項潔，提及該中心一向希望透過資訊技術，創造出一個適合人文研究的數位環境。而這個理念，反映在該中心所製作的每一個資料庫中。臺灣大學歷史系教授吳密察則呼應祝平次教授所說，強調人文學者應該找到自己的位置。他認為過去的合作模式，往往由資訊學者主導，現在應該是個好的時機，反思過去不足的部份。臺灣大學歷史系教授周伯勳，亦強調研究工具增加的同時，扎實的人文學訓練更有其價值。其他來賓則針對教學和推廣，或影視史學等問題，發表各自的意見。在場也有不少新一代的研究生，從學生的角度分享經驗。

下午的場次，則是四場教學的課程。由於篇幅有限，在此無法一一詳述細節。但關於課程規劃，主要基於兩個想法。第一，這些課程需要提供與會者具體操作的經驗；但更重要的是，透過以具體的資料與實際

的案例作為教學內容，與會者不單純是學習工具的使用，更能了解技術可以如何與人文研究接軌、技術可能在哪些層面上幫助人文研究。唯有如此，研究者才有可能更加自由地運用各種技術，能夠想像在數位的時代下「什麼是可能的」。這一個層面，過去類似的教學課程似乎較少著墨，卻是我們想要強調的。這也是工作坊的名稱中會有「想像力」的原因。

除了此次安排的四項工具教學之外，其實還有兩項受到許多人文學者關注的技術工具，由於時間及師資不足，未能安插於課程中，在此稍作介紹，供有興趣的研究者了解。一是社會網絡分析（social network analysis），這以往是社會學者用來觀察大型人際網絡的技術，但近幾年，歷史或文學領域中，有越來越多的學者想利用社會網絡分析工具幫助呈現或釐清文本中複雜的人物關係，或觀察文獻引用的關係，現成的工具有 Gephi [<https://gephi.org/>]。二是 topic modeling，這是資訊學家發展出的一種文本分析技術，能幫人找出難以一一閱讀分類的大量資料中隱含的各類主題，有興趣者可以參考這個由 MITH 舉辦的 topic modeling 工作坊：<http://mith.umd.edu/topicmodeling/speaker-videos/>。

臺灣過去數位典藏的成果豐碩，數位化的大部分文物，都屬於人文研究的領域，對於數位人文的發展，已經扎下良好根基。然而，要怎麼從資料的「保存」轉到「使

用」，應該是需要進一步思考的重要問題。就工具面而言，國內外許多機構都致力於建造多數人文學者能用到的工具。如 CHNM 開發的 Zotero 書目管理工具（免費）[<http://www.zotero.org/>]，及 Omeka 數位物件的展示平台 [<http://omeka.org/>]，Voyant 文本分析工具 [<http://voyant-tools.org/>]，WorldMap 網路 GIS 分享工具等 [<http://worldmap.harvard.edu/>]。這些工具大多不需要艱深的技術背景，強調具有高度彈性，且易於上手，因此有了更加廣泛的應用層面。

就資料和技術兩方面而言，臺灣在數位人文領域中都有其優勢。但我們仍須注意兩點。第一，目前許多資料仍只是鎖在資料庫中，可供查索，卻很難與資料分析工具結合，從不同的角度分析，發揮更大的價值。第二，在技術方面，許多現有的工具是在歐美世界發展出來的，未必能直接應用在其他語文（如中文）之上。要解決這些障礙，前者有賴國內資料持有者更加開放的態度，或建立完整的制度，讓有意願嘗試大量資料（big data）分析的學者，有正常管道可循。後者則有待資訊學者的努力。同時，類似 THC21-Taiwan 的工作坊，應該也能讓人文研究者更為熟悉現有的工具，進一步弭平數位落差。如此，人文研究者也可擺脫對於

技術人員的過分依賴，而可能發展出以自身為主體的數位人文研究。我們期待未來在以上三個方面都能見到積極的發展，想必在不久的未來，臺灣的數位人文會有不同的一番風景。

【註釋】

1. 數位人文目前尚未有統一的定義。我們在另一篇文章曾這樣界定：它指的是那些唯有借助數位科技方能進行的人文研究。反過來講，數位人文的研究，即是企圖尋找在前數位時代中難以觀察的現象、無法想像的議題與無法進行的研究。見項潔、涂豐恩，〈導論——什麼是數位人文〉，收於項潔編，王泰升等著，《從保存到創造：開啟數位人文研究》（臺北：國立臺灣大學出版中心，2011年），頁9-28。另見項潔編，金觀濤等著，《關於數位人文的思考：理論與方法》（臺北：國立臺灣大學出版中心，2011年）；林滿紅等著，《數位人文在歷史學研究的應用》（臺北：國立臺灣大學出版中心，2011年）；劉昭麟等著，《數位人文要義：尋找類型與軌跡》（臺北：國立臺灣大學出版中心，2012年）。
2. 請見會議網站 http://thc21tw.blogspot.com/p/blog-page_4.html。
3. 關於 TEI 的用途，請參考祝平次教授在本工作坊中的兩場演講內容，請見工作坊網站上的錄影。
4. 譬如 Rice University 剛於 2013 年 4 月舉辦的 Digitization in the Humanities 工作坊：<http://hrc.rice.edu/digitizationhuma/>。