

# 數字人文的價值與悖論

張耀銘

[提 要] 數字人文的概念和詮釋林林總總,並未形成統一的共識。在數字技術與傳統人文學科跨界、融合研究的過程中,學術團體、學術會議和學術期刊構成了組合性的學術共同體,並推動了數字人文的興起與發展。數字人文應用於傳統人文學科研究的價值,主要體現在四個方面:一是從文獻檢索數據庫到量化分析數據庫的飛躍;二是從地圖到地理信息系統和空間分析的飛躍;三是從數據化到文本挖掘研究的飛躍;四是從文本到圖形圖像可視化的飛躍。但數字人文應用於人文學科研究,也存在一些悖論:一是數字人文研究的大多數項目還沒有產生全新的敘述,更沒有出現震動學界的成果;二是某些數字人文倡導者推崇數據中心論,從而使人文研究失去人文意義,削弱甚至消解了研究者的主體性;三是在“一切皆可量化”的理想與數據獲取現實之間存在較大差距。我們在享受數字人文帶來的各種便利時,也要避免成為“計算”和“數字”的附庸。

[關鍵詞] 數字人文 人文學科 學術期刊 價值 悖論

[中圖分類號] C39 [文獻標識碼] A [文章編號] 0874-1824(2019)04-0105-14

近年來,數字人文已經成為學界、業界的熱點,儼然有顯學之勢。2011 年武漢大學數字人文研究中心成立,是中國大陸首個數字人文研究中心。此後,中國人民大學信息資源管理學院成立數字人文技術實驗室,北京大學成立數字人文建設與發展研究課題組,曲阜師範大學成立數字人文研究中心;南京大學更是雙管齊下,歷史學院成立數字人文研究中心,人文社會科學高級研究院成立數字人文創研中心。隨著研究機構的成立,各種學術會議相繼召開,如北京大學主辦的三屆數字人文論壇,主題分別是“跨界與融合:全球視野下的數字人文”、“互動與共生:數字人文與史學研究”、“孵化與實踐:需求驅動下的數字人文項目”;南京大學主辦的“數字人文:大數據時代學術前沿與探索”、“比較視野中的數字人文反思”;全國高等學校文科學報研究會主辦、《河南大學學報》編輯部與河南大學文學院承辦的“數字人文時代的文學研究”學術研討會等等,暢抒己見,熱烈爭鳴,令人目不暇接。本文將圍繞數字人文的發展源流、學術期刊在數字人文興起中發揮的重要作用、數字人文應用於傳統人文學科研究的價值與悖論,談點不成熟的意見,求教於方家。

## 一、“數字人文”概念及其發展源流

我們正進入一個由數字技術和互聯網造就的萬物皆媒的全新時代,各種新興突破性技術集中

出現,正以日新月異的驚人速度和不可限量的變革潛能在不斷推出其最新形態。各行各業都在發生重大轉變:舊的商業模式被顛覆,新的商業模式在出現;教育、金融和交通體系被重塑,生產方式與生活方式被刷新;技術革命與知識更迭的週期不斷縮短,新的學術方法和理論取代舊的學術方法和理論也將成為常態。

### (一)“數字人文”的概念形成

數字人文(Digital Humanities,簡稱DH)這一術語,20世紀末美國弗吉尼亞大學人文科學高級研究所等機構就已經開始使用。2000年,斯坦福大學英文系教授弗蘭克·莫雷蒂(Franco Moretti)在後來成為經典的《世界文學的猜想》中提出一項研究“世界文學”的方法,就是遠離“細讀”傳統,發展出一種新的“遠讀”模式,通過量化的方法,對龐大的文本體系中的類別因素和形式元素作出解釋。“遠讀”概念的提出,預言了以計算機和“文本”結合來考察文學體系的做法。這被認為是“數字人文”文學研究在21世紀發生發展的一個轉捩點。莫雷蒂也由此被認為是數字人文的代表性人物。大約在2001年,英國布萊克威爾出版社編輯安德魯·麥克內利約請弗吉尼亞大學英文系教授兼大學圖書館主任約翰·安斯沃思與蘇珊·施賴布曼和雷蒙德·西門斯三位領域巨擘討論出版物《人文計算年鑒》的更名問題。<sup>①</sup>雷蒙德·西門斯主張用當時比較流行的術語《人文學計算指南》作為書名;布萊克威爾出版社的編輯與銷售人員從市場營銷角度出發,提出一個更吸引讀者眼球的書名《數字化的人文研究指南》;約翰·安斯沃思反對書名出現“簡單的數字化”,建議使用《數字人文指南》。<sup>②</sup>各方堅持己見,難以統一,更名的事情因此被擱置。不過“數字人文”這個新生事物,卻得到眾多學術會議、學術期刊和社會人士的支持。<sup>③</sup>2001年,美國加州大學洛杉磯分校信息研究學院教授約翰娜·德魯克在ACH/ALLC會議上做主題發言,標題即為《現狀核實:數字人文項目與前景》。從2002年開始,“數字人文”概念的使用得到了大幅度的提升。

關於“數字人文”的概念和詮釋林林總總,並未形成統一的共識。梳理目前的文獻資料,大體上有四種不同的代表性觀點。第一,把“數字人文”理解為一種實踐。約翰·安斯沃思教授定義為“一種代表性的實踐,一種模擬方式,一種推理,一種實踐本體核心的思路。這種代表性的實踐可分為兩個方面,一方面是高效的計算,另一方面是人文與科技的溝通”。<sup>④</sup>第二,把“數字人文”看作一種研究方法。“它是人文學科研究方法的補充,這一方法基於對計算機和互聯網的運用,是一種更廣闊意義上的‘計算’。對提出的問題採用不同的研究方法並研究一些目前尚未解決的問題,這構成了‘數字人文’。”<sup>⑤</sup>第三,把“數字人文”視為一個研究領域。數字人文不僅發展迅捷,而且更具擴張性。它除了歷史學、文學、社會學、語言學、考古與藝術學之外,有時也包括“平台研究”、“軟件研究”、“媒介研究”,以及像文本編碼、數據挖掘和數據分析等新領域。<sup>⑥</sup>第四,把“數字人文”看成為一門學科。英國劍橋大學教授凱特·巴雷特在《從小冊子到像素:轉變中的人文學科》一文中提出了“數字人文學科”概念,是指用電子技術來處理和分析傳統人文學科研究資料,並結合人文研究方法論和電子工具創建電子歷史文本的人文研究。特別有趣的是巴雷特還定義了“數字人文學者”的形象:他們不僅能在計算機實驗室中操作數據庫和編程軟件,而且能夠跨越狹義人文學科定義,進入與電子技術結合的廣闊人文學科研究平台。<sup>⑦</sup>歐美許多大學已經將“數字人文學科”納入正式教育體系,涉及語言學與語言研究、計算機科學、文學和語言研究、歷史、圖書館學與信息科學、人類語言技術等學科,開設“數字文化”、“數字人文”、“大數據與社會化、文化”、“數字內容管理”、“數字人文研究”、“數字資產媒介管理”等系列課程,包括了本科、碩士、博士的整個階段。

### (二)“數字人文”的發展源流

“數字人文”概念的提出,其實才有十幾年的時間。但要追根溯源到“人文計算”,已經有 70 年的歷史。人文計算是通過引進計算機、統計學等技術與方法對傳統人文學科(例如歷史學、哲學、語言學、文學、藝術、考古學、音樂等)的問題進行輔助性研究的一種實踐。英國倫敦藝術大學教授蘇珊·霍基在《人文計算的歷史》一文中,對“人文計算”的歷史進行了整體性梳理,將其分為“起始”、“聯合”、“新發展”和“成熟”四個階段。

### 1. “人文計算”的起始階段(1949~1970 年代)

“人文計算”(Humanities Computing),可以追溯到義大利耶穌會羅伯托·布薩(Roberto Busa)神父與 IBM 合作,用穿孔卡和磁帶存儲的計算機,輔助編制義大利著名神學家、哲學家托馬斯·阿奎那及相關作者作品中 1,100 萬中世紀拉丁文詞彙的索引。工作從 1949 年開始,先後在義大利米蘭、皮薩、威尼斯和美國的科羅拉多州博爾德等地進行,團隊最多時達 66 人。經過近 30 年的不懈工作,至 1980 年編成了 56 卷百科全書式的《托馬斯著作索引:聖托馬斯·阿奎那全集索引和語詞索引》,由德國斯圖加特市弗羅曼—霍爾茲柏格出版社印行。2005 年,已經 92 歲高齡的布薩與義大利“托馬斯·阿奎那基金會”合作,把托馬斯著作索引數據庫推上萬維網,使其內容更翔實,檢索更方便。布薩的貢獻主要體現在三個方面:一是運用計算機將文本轉化為數據的方法,具有革命性的開拓作用;二是把索引條目詞目化,從而推動了詞目化軟件的開發,使詞典編纂從中受益;三是對托馬斯著作的文獻數據處理程序,經過必要的技術改進,“也會適用於法律、化學、醫學、工程技術等方面的文獻分析和信息檢索”。<sup>⑧</sup>布薩由此成為把計算機應用於人文科學的先驅者和“人文計算”的奠基人。

### 2. “人文計算”的聯合階段(1970 年代~1980 年代中期)

這個階段,學者思想開放,學術會議增多,學術交流活躍。英國開啟的“雙年系列會議”成為人文學科計算的重要焦點,北美啟動的“人文學科計算國際研討會”系列會議影響提升。在此基礎上,“文學與語言學計算學會”(ALLC)在英國成立,“計算機與人文學科學會”(ACH)在北美成立。這個階段的“聯合”,主要體現在兩個方面:一是學術會議促進了學術共同體的建構,學術共同體反過來又推動了學者群體的聚合、學術期刊的創辦和學術論文的生產。二是人文學者和計算機技術人員之間有了聯合,從而開發出了牛津語彙索引程序(OCP)和希臘語庫(TLG)為代表的一系列軟件程序,大大減少了人文項目在程序設計方面的費用,並促進了學科之間的交叉融合。

### 3. “人文計算”的新發展階段(1980 年代中期~1990 年代早期)

這個階段發生的三個事件與人文計算的發展有著密切的關係,值得一提。(1)個人計算機的出現。計算機最早應用在二次世界大戰中的軍事領域,1950 年代計算機應用在大學裡快速增長,1960 年代計算機支撐的研究方法在社會科學領域得到應用,1970 年代隨著數據庫管理系統、網絡管理系統的出現,應用領域從科學計算、事務管理逐步走向人文研究領域。1981 年 8 月 12 日,IBM 公司正式推出了全球第一台個人計算機——IBM PC,它首創了“個人計算機”的概念,並為 PC 制訂了全球通用的工業標準。這一具有劃時代意義的發明,從此掀開了個人計算機普及的序幕。隨著越來越多的紙本圖書與期刊被數字化,並從圖像掃描轉換為文本字符,也促使更多學者依靠快速進步的計算機語義識別技術處理更大量的文本數據。(2)電子郵件的流行。雖然電子郵件是在 1970 年代發明的,但卻是在 1980 年代才得以流行。隨著個人計算機的興起,電子郵件開始在學者以及大學生中廣泛傳播開來。當時的大多數網絡運營商提供了跨國收發電子郵件的服務設施,並創建了郵件自動回應程序軟件。特別是 1987 年郵寄清單 Humanist 的創建,為人文計算領域學術

共同體提供了一個定期保持聯繫和學術交流的社區平台,其意義非同小可。(3)文本編碼倡議(TEI)的貢獻。TEI起源於“文本編碼創始項目”,用大約400個編碼標記了一個結構,可以輕而易舉地向新的應用領域進行擴展。該結構由兩部分構成:一是各種類型文獻的標識,二是此文獻的書目信息(元數據部分)。<sup>9</sup>1994年5月,TEI以紙媒形式和電子版形式發佈《電子文本編碼和交換指南》。這被認為是“歷史上第一次系統化地對所有人文學科文本進行分類和定義”。<sup>10</sup>

#### 4. “人文計算”的成熟階段(1990年代早期~2003年)

這個階段的所謂“成熟”主要體現在兩個方面:(1)互聯網的廣泛應用。經過二十多年的高速發展,1991年第一個網頁和基於文件內容的搜索協定誕生,1993年第一個圖形瀏覽器出現,1994年電子商務現身網絡,1995年亞馬遜推出在線購物服務,1996年由微軟提供的免費電子郵件服務啟動,1997年博客讓互聯網成為私人交談空間,1998年谷歌上線標誌著網上搜索時代來臨,2003年網絡電話VOIP成為主流,社交網絡聚友網發佈。隨著計算機的多媒體化及其與互聯網的結合,“計算機不僅能處理文本對象,也可以處理圖像、聲音,同時用大規模數據庫來存儲和管理這些多媒體信息,更多的語料庫、聲音和圖像數據庫被建立”。<sup>11</sup>歷史文獻數據和新近產生的研究文獻大部分被數字化,整個人文計算研究的基礎也發生了革命性的變化。(2)“檔案館”項目密集出現。這個時期由美國國會圖書館和美國各大學等機構資助的項目增多,有相當數量的項目與檔案或歷史記憶主題相關,如美國弗吉尼亞大學數字歷史研究中心1993年啟動的以美國南北戰爭時期平民生活檔案為主題研究的“影谷檔案”,美國弗吉尼亞大學人文高新技術研究所2001年啟動的以美國傳奇詩人Emily Dickinson個人檔案為主題的“狄金森電子檔案項目”,美國喬治梅森大學歷史與新媒體研究中心2002年啟動的“911事件數字檔案項目”,美國弗吉尼亞大學人文高新技術研究所2003年啟動的以阿根廷查科省歷史檔案為主題的“查科研究檔案項目”。<sup>12</sup>此外,還有美國女作家檔案、羅塞蒂檔案、威廉·布萊克檔案等等。根據國外檔案數字人文項目劃分標準,國外檔案數字人文項目主要包括五大類:GIS歷史地理可視化項目、語料庫建設項目、歷史數據庫建設項目、社會及歷史場景重建項目、檔案數字化項目。“數字人文研究與檔案、檔案工作緊密相連,目前,全世界的檔案學者都已對檔案在數字人文中的角色、功能和參與實踐活動等進行探討和研究。”<sup>13</sup>

#### 5. “數字人文”階段(2004~)

1930年,陳寅恪在《陳垣〈敦煌劫餘錄〉序》中指出:“一時代之學術,必有其新材料與新問題。取用此材料,以研求問題,則為此時代學術之新潮流。”<sup>14</sup>通過數字人文提供的新方法和新實踐,建立數據庫、發掘新材料,驗證舊結論、研究新問題,正成為我們這個時代的新潮流。

##### (1)《數字人文指南》出版

經過幾年沉澱,安斯沃思命名的《數字人文指南》在2004年出版。這標誌著人文主義與技術主義的爭論暫時告一段落。“但名稱的轉換並不僅僅是為了吸引更多人文學者加入該領域而採取的策略,它還是從內部技術的發展和學科外延的擴大而對學科內涵和外延的一種修正。”<sup>15</sup>儘管該書中“人文計算”出現的頻率次數是“數字人文”的大約兩倍,但對“人文計算”進行總結的同時,也對“數字人文”進行了推廣和前瞻,這對數字人文以後的研究與發展無疑具有里程碑的重要意義。從此,“數字人文”逐漸取代“人文計算”,成為計算機協同人文學科開展合作性研究的代名詞。

2004年起,“歐洲研究基礎建設策略論壇”聯合英國、法國、德國、荷蘭、丹麥等國的研究機構和力量,合作推動“藝術與人文的數字研究基礎建設”。2008年,美國國家人文基金會設立數字人文辦公室並啟動“數字人文進步獎”,申請的項目內容包括:一是運用數字實驗和計算等技術對人文

學科開展創新研究,增強實踐能力;二是分析數字文化對人類社會生活造成的影響;三是探索數字人文學科在特定領域或學科研究中的意義和作用,以及具有的哲學影響;四是振興與恢復已損壞但依然有利用價值的數字項目,這些項目對人文學科的科研和教學有實質性幫助,並承諾可以提高公眾的人文素養。<sup>⑩</sup>這成為數字人文領域被制度化的重要標誌。此後,數字人文技術創新充滿了張力,許多人正在這個領域做著充滿想像力的探索。“這種張力存在於那些已經進入這領域很長時間的人和那些正準備進入的人之間,存在於學科性和跨學科性之間,存在於製作和解讀之間,存在於這個領域的歷史和未來之間。”<sup>⑪</sup>

### (2) “數字人文”的技術體系

安妮·伯迪克、約翰娜·德魯克等在《數字人文:改變知識創新與分享的遊戲規則》中這樣描述,當代數字人文尋求超越簡單的數據倉儲建設和編輯從而走向新的整合性實踐,它包括開展統計處理(計算機語言學)、運用鏈接(超文本)、建模(包含結構構建與視覺呈現)、創造結構化數據(可擴張標記語言)、開展循環反復編輯並採用版本控制等。劉煒等認為數字人文的技術體系主要包括六個方面:一是數字化技術:構造數字人文的資源基礎;二是數據管理技術:構造數字人文服務系統;三是數據分析技術:構造數字人文應用平台;四是可視化技術:構造數字人文的直觀形象;五是VR/AR技術:構造數字人文的交互環境;六是機器學習技術:構造智能系統實現智慧服務。<sup>⑫</sup>

總之,數字技術工具不僅是承載和處理人文科學內容的“形式”,其本身就承擔了知識界面、3D建模、文本編碼、數據庫設計、地理空間分析、歷史仿真、個性化服務與精準推送等實質性功能,可以引導、擴展甚至改變學者的研究習慣。當前由人工智能引領的新一輪科技革命方興未艾。在大數據、數字人文、超級算法等新理論新技術的驅動下,古老的人文學科正呈現智能檢索、數據分析、內容挖掘、場景模擬、跨界融合、平台協作、機器學習等新特徵,這將促使傳統的人文學科大量借用自然科學的技術手段進行研究,“科學性”顯著增強。

### (3) 高校數字人文教育

近十年來,各個大學相繼組建數字人文實驗室、招收數字人文專業的研究生已經成為普遍現象。據國際數字人文組織聯盟(ADHO)的統計,截止到2018年3月,全球有194個冠以“數字人文”的機構或項目正在運行。<sup>⑬</sup>數字人文中心建設起源於美國,從全球的分佈情況來看,絕大多數的高校數字人文中心集中在北美和歐洲國家,亞洲國家只有少數的專門研究機構。美國知名的高校數字人文中心,主要有弗吉尼亞大學人文先進科技研究所、內布拉斯加大學林肯分校數字人文研究中心、斯坦福大學數字人文實驗室、加州大學洛杉磯分校數字人文中心、哈佛大學“人文2.0實驗室”、伊利諾伊大學香檳分校人文藝術與社會科學計算研究所等。此外,英國倫敦大學學院數字人文中心、英國倫敦國王學院人文計算中心、英國格拉斯哥大學人文科技信息研究所、英國謝菲爾德大學數字藝術與人文專門研究中心、德國哥廷根大學數字人文中心、加拿大維多利亞大學人文計算與媒體中心、日本立命館大學日本文化藝術數字人文中心、中國武漢大學數字人文研究中心和中國台灣大學數字人文研究中心等,也取得不俗的研究成果。這些數字人文中心建設模式體現了四個方面的特徵:一是數字人文中心建設主要是由綜合性大學推動的;二是絕大多數定位於“跨學科協同創新服務機構”;三是“項目制”是數字人文中心運行的基本方式;四是通過開展各種形式的活動促進信息技術和人文社會科學的融合。<sup>⑭</sup>

數字人文教育在歐美國家起步較早,且發展迅速。早在1995年,加拿大麥克馬斯特大學就開始為人文學科的學生開設“人文計算引論”、“人文領域的多媒體介紹”、“電子文本和計算語言學”

課程,旨在為學生提供信息技術基本方法。2002年英國倫敦大學國王學院數字人文系正式成立,2005年倫敦國王學院率先成立了數字人文博士學位授予點。隨後的十幾年中,美國的馬里蘭大學、德國的帕紹大學、義大利的博洛尼亞大學、瑞士的洛桑大學、愛爾蘭的科克大學、荷蘭的阿姆斯特丹大學、加拿大的多倫多大學等幾十所大學開設了數字人文課程。2018年,美國哈佛大學東亞語言與文明系開設了《東亞數字人文課程》,課程內容包括文本數據分析與處理、數據可視化、數據庫技術、地理信息系統等。據歐洲語言資源與技術研究基礎設施及藝術與人文數字研究基礎設施合作建立的數字人文課程註冊門戶網站顯示,截止2018年2月,在網站中註冊的數字人文課程共有153個。這些課程包括了本科、碩士、博士的整個階段。其中面向本科生的有39個,面向碩士生的有69個,面向博士生的有5個。此外,還開設了28個學分課程,4個暑期學校課程,1個繼續教育課程,4個模塊課程。<sup>21</sup>美國與歐洲國家的數字人文課程涉及的學科種類較多,跨學科特徵明顯。學生除了學習專業課程之外,還在課堂上針對數字人文中遇到的實際技術問題,結合相關實用的工具類軟件進行練習操作。<sup>22</sup>

## 二、學術期刊在數字人文興起中發揮的重要作用

綜觀數字人文的歷史,從布薩神父與IBM在1949年的合作開始,經歷人文計算的聯合、新發展、成熟到數字人文,數字技術與人文研究的跨界、融合已經形成了清晰的發展脈絡。其中,學術團體、學術會議和學術期刊,構成了組合性的學術共同體,並推動了數字人文的興起與發展。

### (一)學術共同體成為學術秩序的建設者、維護者

20世紀40年代,英國哲學家邁克爾·波蘭尼(M. Polanyi)在文章《科學的自治》中首次提出“學術共同體”這個概念。他把具有共同信念、共同價值、共同規範的從事科學研究的科學家稱為學術共同體,以區別於其他社會群體與組織。美國科學史家托馬斯·庫恩指出:“科學共同體是由一些科學專業的實際工作者組成。他們由他們所受教育和見習訓練中的共同因素結合在一起,他們自認為也被認為專門探索一些共同的目標,也包括培養自己的接班人。”<sup>23</sup>學術共同體是什麼?中國科學院院士韓啟德認為,所謂的學術共同體就是指一群志同道合的學者遵守共同的道德規範,相互尊重、相互聯繫、相互影響,共同推動學術的發展,由此而形成的群體。<sup>24</sup>由此可見,學術共同體是基於某種學科、價值、理念和規範的認同而形成的結構鬆散的學者社群與學術團體。

從“人文計算”到“數字人文”的轉變,使該研究領域逐漸獲得世界人文學界的承認,從中我們看到學術共同體發揮的作用相當清晰。自1973年開始,在歐美先繼成立了一些學術團體,比較重要的有文學與語言學計算學會(ALLC)、計算機與人文學科學會(ACH)、美國人文計算學會、澳大利亞數字人文協會、日本數字人文協會、歐洲數字人文協會和國際數字人文組織聯盟(簡稱ADHO)等。這些學術共同體由學者以專業為基礎自主結社和自主運作,制定規則、召集會議、創辦期刊、同行評審、學術培訓、發現人才,學術共同體通過學術交流、學術批評、學術評論表達“民意”和進行監督。可以說,“這種由團體、刊物和會議所構成的組合性學術共同體,不啻是一個開放的學術討論和批評的空間,是學術秩序的建設者、維護者和監督者”。<sup>25</sup>

### (二)學術會議成為數字人文討論和批評的開放空間

自從1970年在英國劍橋召開第一次ALLC人文計算會議之後,逐漸形成通過會議定期開展學術交流與批評活動的機制,以後更發展成為在歐洲和北美輪流召開的慣例。20世紀70年代,ALLC會議召開了四屆,ICCH會議召開了三屆。2016年開始,由密歇根大學數字人文項目發起的

年度數字人文研討會(簡稱 GDHS)已舉辦三屆。1977年8月,第三屆人文領域計算機國際會議由加拿大蒙特利爾大學和滑鐵盧大學共同主辦,參加會議的有人文、計算機、數學、工程等領域的專家,跨學科交流了詞典學、歷史、音樂、舞蹈、文學數據管理與計算機輔助教學等領域的新研究發現,並出版了論文集。從此以後,編輯出版會議論文集,成為擴大數字人文會議影響的重要形式。有學者分析了近50年國際數字人文研究的高被引文獻,在排名前十(被引頻次 $\geq 10$ )的知識來源中,圖書類文獻8部,學術論文2篇,表明圖書對推動數字人文知識結累和研究演化的貢獻力度更大。明尼蘇達大學出版社2012年出版的《數字人文學科的爭論》排名第一。這本會議論文集,收錄美國數字人文教育機構、高校以及圖書館領域的數十位專家發表的論文,“旨在探索數字人文理論、方法和實踐,試圖挖掘出數字人文的發展潛力與張力”。<sup>⑤</sup>麻省理工學院出版社2012年出版的《數字人文研究》,排名第二。該報告不僅回答了“數字人文是什麼?”這一核心問題,而且提供了一個考察數字人文領域的深邃視角。

在美國學術界,學術會議自始至終都是一個學術評價的過程。1999年和2002年,美國弗吉尼亞大學分別舉辦了“人文學計算是大學學科嗎?”和“數字人文課程研討班”兩個重要的學術研討會。會議由英文系教授傑羅米·麥根、約翰·安斯沃思和加州大學洛杉磯分校信息研究系教授約翰娜·德魯克發起並組織,會前對申請參會的論文進行評審,論文質量合格的作者才有資格參會。會議中的報告、評議、問答等流程,既是對提交論文進行的學術評價,也是學術爭鳴、學術批評的重要環節。來參加這兩次研討會的史蒂文·拉姆塞、馬修·克申鮑姆、本特尼·諾維斯基等人,後來都成為數字人文研究領域最重要的代表人物。兩次會議主題雖各有側重,但又能看出其傳承與發展的脈絡,會議最終“正式將人文計算塑造為數字人文,並將目錄校勘等有關文本等學術工作定義為它主要的內容”。<sup>⑥</sup>2004年,約翰·安斯沃思與S.Schreibman、R.Siemence聯合在布萊克威爾出版公司出版《相伴數字人文》(*A Companion to Digital Humanities*),該論文集收錄37篇原創性成果,提供了一個全面而簡潔的數字人文研究評述,書中明確提出要克服人文計算“單純數字化”的傾向,從技術視角回歸到研究對象相關知識體系的數字化呈現上來,建立“數字人文”新視角。<sup>⑦</sup>2006年,許多原本屬於人文計算領域的會議與研究中心紛紛更名搭上“數字人文”的快車。2007年,美國馬里蘭大學創建“數字人文中心網絡”,依託關係網絡和資源平台更加便利地與國際同行在項目共建、工具共用和知識共享等領域開展合作。目前已有兩百多家一流高校的學術機構加入,成為重要的數字人文研究與協作的國際性科研組織。2009年,北美人文學科領域最重要的現代語言協會年會召開,會議上諸多討論的主題都是圍繞著數字人文展開,諸如何種研究可以算作數字人文研究,怎樣的學者算作領域內的研究者,數字人文本身的目的到底是完成項目、探究理論化、研究計算技術、跨學科溝通、實踐技術應用,還是單純的科研經費需要,等等。<sup>⑧</sup>學者們對數字人文本身存在很多不同理解、不同見解,交流、爭論、批評成為常態,這既造就了開放的學術討論和批評的空間,也彰顯了數字人文獨特的學術魅力。

### (三) 學術期刊成為傳播數字人文的重要平台

人文計算的實踐公認以1949年羅伯特·布薩神父為托馬斯·阿奎那及相關作者的著作編制語詞索引為開端。1966年由約瑟夫·雷賓主編的《計算機與人文學科》雜誌創刊發行,標誌著作為學科領域的人文計算正式走向學術舞台。<sup>⑨</sup>1986年由ALLC在牛津大學出版社出版的《文學與語言計算》,被認為是數字人文領域最早的學術期刊。隨著時代的發展,這本期刊也與時俱進,先改名《文學與語言計算:人文領域的數字學術》,後更名《人文領域的數字學術》,而被業界廣泛熟知。<sup>⑩</sup>

2007年出版的《數字人文季刊》，直接拋棄傳統的紙本模式，採取在線出版，涵蓋人文學科數字媒體的各個方面。國際數字人文組織聯盟，目前擁有《文學和語言計算》、《文本技術》、《人文領域計算機應用工作論文》、《數字人文季刊》四本同行評審期刊，通過召集圓桌會議、發表專題文章、組織書評和“回顧性”評論，向全球傳播數字人文研究的理念、方法和成果。這些學術期刊，已經成為數字人文學術傳播穩定的基本單元，成為數字人文學術共同體成員聚合、交流、監督的重要中心。“如果沒有學術期刊作為組織的公共平台，那麼，共同體就難以聚合，其邊界也難以清晰，學術交流則難以高效地展開，共同體在研究和評價中的作用勢必難以發揮。”<sup>②</sup>

學術共同體的形成取決於諸多條件，其中學術期刊對於學術共同體的產生、發展有著不可替代的作用。首先，學術期刊劃定了學術共同體的邊界。“學術共同體主要包括學術期刊系統和專業學會系統。這二者是學者進入學術共同體、參與並建立正式學術交流網絡的關鍵平台，在學者和‘門外漢’之間劃定了一條明顯的邊界，因此對學者的學術生涯和學術共同體的運行至關重要。”<sup>③</sup>其次，學術期刊成為學術評價的平台。由於學術期刊的辦刊主體是學術共同體成員，所以其除了展示、交流和傳播作者的學術發現和見解之外，還有一個重要屬性就是學術評價。學術共同體是構建學術傳播秩序的當然主角，它可以自主地制定評價規則、評價標準和評價程序，並且主導評價過程。在對成果本身的評價中，學術共同體採用的是複合評價的制度，即學術會議的評議和討論、發表前的評審（包括第三方評議）、論文發表或著作出版後的評價（學術期刊組織專家發表專題文章進行評議）、組織同行專家學術評獎。在這過程中，學術期刊不僅發揮了重要的引領作用，更體現出對學術標準、學術方法、學術品質、學術評價和學術精神等方面的公信力和權威性。再次，學術期刊成為凝聚學術共同體成員的中心。學術期刊作為學術傳播的中心所在，其作用的發揮離不開學術共同體的認可和支持。所以說，“一個有生命力的學術期刊背後一定站著一個思想活躍的學術團體，而這個學術團體應該是一個世界性的學術共同體的有機組成部分”。<sup>④</sup>以美國歷史學會為例。在學術共同體的努力下，2006年以後的學會年會上，不定期出現數字史學的討論版塊，2014年始更設置了數字史學工作坊，提供各種層次的技能培訓與資源分享。《美國歷史雜誌》從2013年開始闢專欄“數字史學評論”。近六年來，該專欄發表了評論文章101篇，先後點評超過100多個數字史學的研究和項目，突顯了數字史學在美國史學界是一個充滿活力和希望的話題。<sup>⑤</sup>

### 三、數字人文助力人文學科“輪廓重繪”

美國著名數字人文專家約翰·安斯沃思用七個“學術原語”概括數字人文常用的基本方法，即“發現”、“注釋”、“比較”、“參考”、“抽樣”、“說明”、“表現”。該理論後來被歐洲學者布蘭克·托拜西和安德森·希拉等將其規範為五個基本原語，即“發現”、“收集”、“比較”、“傳遞”和“協作”。<sup>⑥</sup>這種數字人文研究的“方法論共同體”，對於傳統人文學科的價值在於，借助數字技術所擁有的多樣性和擴張性，通過五個基本“學術原語”等手段進行跨界研究，從而擴展疆域和潛力，使學術領域實現“輪廓重繪”。對於數字技術的價值在於，“在注入人文主義價值觀和方法後，其功能更加多元和強大，更符合人文學術的本質特徵”。<sup>⑦</sup>

#### （一）從文獻檢索數據庫到量化分析數據庫的飛躍

中國歷史悠久，現存的古代文獻資料浩如煙海。經過20多年的努力，歷史文獻的數字化已經取得了很大的進展，建成了一些可供檢索的文獻數據庫。第一，圖書館、檔案館的古籍、檔案數字化目錄。第二，借助掃描、數碼拍照等技術手段，將紙本轉化為數碼圖像文件，形成了對古籍文獻、檔



案資料、報紙期刊的圖像文件進行存儲和檢索的圖像數據庫。第三,通過 OCR 技術與人工核對的結合,建成全文檢索資料庫。第四,事實型工具書被做成數據庫、軟件或插件。<sup>③</sup>這些文獻數據庫,既有圖書館、檔案館、學術期刊自建的,也有商業公司生產的數據庫產品。如北京大學的《中國歷代典籍總目》、《中國基本古籍庫》,首都師範大學電子文獻國學傳播中心的《古籍電子定本工程》,浙江大學與哥倫比亞大學合作開發的“甲骨文數字化項目”,國家圖書館的“華夏記憶項目”,敦煌研究院的“數字敦煌”,上海圖書館的“民國時期期刊全文數據庫(1911~1949)”,南京圖書館的“中國近代文獻圖像數據庫”,商務印書館的“《東方雜誌》期刊全文檢索數據庫”,瀚堂公司的“瀚堂近代報刊數據庫”等為代表的古籍與專題數據庫的大量出現,為學者們獲得文獻文本並儲存提供了極大的便利。但這些文獻數據庫還處於比較原始的紙質替代狀態,普遍只有檢索功能,並且只能按原始資料的結構進行瀏覽,不能幫助研究者統計、分析文本。

所謂量化數據庫,“統指各種搜尋能夠涵蓋一定地域範圍、具有一定時間跨度的整體性大規模個人或其他微觀層面信息的系統(一手)資料,並將這些資料按照一定數據格式進行電子化,構建成適用於統計分析軟件的量化數據庫並進行定量研究的方法”。<sup>④</sup>美國學術界利用歷史資料進行量化數據庫構建與研究的實踐遠遠超前,其中學者們利用國際微觀共享整合數據庫、美國猶他人口數據庫研究出的很多成果與著作,產生了重大的學術和社會影響,也成為中國學術界構建量化數據庫的參照座標。從 1980 年代起,歷時 20 多年時間,李中清、康文林團隊建立了基於八旗戶口冊和清代皇室族譜資料的中國多代人口數據庫系列(簡稱 CMGPD),包含遼寧、雙城和皇族三個子數據庫。遼寧數據庫涵蓋了 1749~1909 年間遼東地區 26 萬人的 150 余萬條記錄。雙城數據庫涵蓋 1866~1913 年間雙城縣 10 萬人的 130 余萬條記錄,並嘗試與不同時段的家戶地畝資料相連接。這些大規模、長時段的微觀歷史數據包含豐富的人口和社會經濟信息,具有時間上的深度和空間上的廣度,對人口統計學、家與家族、親屬關係、社會分層與流動、衛生健康等多個研究領域有重要價值。<sup>⑤</sup>再比如美國哈佛大學教授包弼德領導的與北京大學中古史研究中心以及“中研院”史語所合作的“中國歷代人物傳記資料庫”(簡稱 CBDB),支持各種各樣的查詢。“包括單個人物查詢、入仕途徑、官職查詢,還可以查詢社會關係網絡、查詢兩個人之間的社會關係,甚至查詢不同地區間人物的關係。”<sup>⑥</sup>所以,包弼德教授又把這種數據庫稱之為“關係數據庫”。量化數據庫的開發與研究對歷史學而言,至少有兩方面非常重要的價值:“一是豐富了歷史學研究的方法論與範式;二是能有效利用大規模材料,實現不同材料的彼此連結,推動大規模、合作化研究形式的出現。”<sup>⑦</sup>

## (二) 從地圖到地理信息系統和空間分析的飛躍

英國地理學家羅格·湯林森最早提出地理信息系統概念,並在 1971 年建立了加拿大地理信息系統。地理信息系統(Geographic Information System,簡稱 GIS),以計算機技術為基礎,通過對地理原始數據進行數據化的處理、存儲、提取與管理,通過數據建模、空間分析、專題製圖等手段,研究其與歷史、社會、自然之間的關係,探索發展演變的規律和趨勢。1990 年代之後,西方發達國家相繼建立了一些頗具特色的歷史地理信息平台,如美國東海岸公共圖書館開發的弗吉尼亞 1870~1935 的鐵路項目、紐約圖書館開發的紐約城市地理信息系統項目、加州大學伯克利分校開發的菲律賓文化地圖、布朗大學開發的非洲史動畫地圖集和瑞典於默真大學開發的拜占庭帝國猶太社區項目、愛爾蘭皇后大學數據數字化與分析中心開展的“愛爾蘭土豆大饑荒研究”等。<sup>⑧</sup>隨著地理信息技術的快速發展,各國的學者們相繼利用這項技術開展空間分析與模型建構,並取得了一系列的科研成果。1992 年,美國地理學家 Michael F. Goodchild 教授正式提出“地理信息科學”概念。從此,

地理信息系統不再僅僅是一門技術,更被廣泛地視為一門科學。

地理信息系統及其相關平台和方法的利用,已成為數字人文中不容忽視的要素和不可或缺的一部分。中國在這方面的起步較晚,最初甚至要借助國外的基金、技術與人才,以尋求發展之路。(1)中國歷史地理信息系統(簡稱 CHGIS)。該項目 2001 年 1 月正式啟動,由羅斯基金會資助,主要參加機構為哈佛大學與復旦大學中國歷史地理研究中心等。CHGIS 項目以秦朝建立的公元前 221 年到清朝滅亡的 1911 年為時間範圍,試圖建立一套中國不同歷史時期政區連續變化的基礎地理信息庫,為研究者提供 GIS 數據平台、時間統計以及查尋工具和模型。通過 CHGIS 系統的分層地圖,可以看到一個縣的轄區是如何與當地的地理景觀產生聯繫的,可以看到歷史信息的空間關係,從而解決了“時間—空間”維度的整合問題。(2)中國歷代人物傳記資料庫(簡稱 CBDB)。CBDB 利用中國歷史地理信息系統的數據,在人物傳記資料的組織與展示上融入“空間分析法”,將歷史人物的籍貫、生平、仕途、遊歷、關係、著作、死亡等置於多維時空架構之中,為研究者提供了全新的方法。例如,從 CBDB 中提取信息並映射到地圖上,即可生成明代進士地理分佈圖。明代進士主要來自江南地區,東南以及江西遠遠超過其他地區。“仔細研究明代進士的地理位置,就會發現明代政治體系的空間特點。”<sup>④</sup>(3)中華文明時空基礎架構(簡稱 CCTS)。該項目由台北“中研院”主持,主要內容包括基礎歷史地理圖資、Web GIS 的整合應用環境以及主體性空間信息三大部分。基礎歷史地理圖資以譚其驤主編《中國歷史地圖集》為主要範本,提供中國從先秦到清朝歷代的疆域,並結合丁文江 1930 年代所編《申報地圖》,構成上下逾兩千年的中國歷代基本歷史地理底圖。以 1990 年代百萬分之一《中國數字地圖》為現代底圖,同時結合歷代黃河變遷、漢代古墓分佈、清代糧價人口、明清江南市鎮等多個圖層,整合台北“中研院”漢籍全文檢索系統、地方誌檢索系統、古今地名檢索系統等,期望透過“時間—空間”的基礎信息架構,提供多學科研究成果,發展出各種專題地理信息系統,進而促成多學科之間的交流。<sup>⑤</sup>此外,“台灣歷史文化地圖系統”、“民國時期北京都市文化歷史地理信息數據庫”、“華夏民族家譜地理信息系統”、“絲綢之路歷史地理信息開放平台”等也已經取得了一定的成績。總之,GIS 技術為現代歷史地理學注入了新的活力,“許多過去無法解決的難題通過 GIS 手段得以解決,它的開發價值也越來越受到學者的重視”。<sup>⑥</sup>

### (三)從數據化到文本挖掘研究的飛躍

數字化、數據化和文本挖掘,是歷史文獻數據庫的三種不同形態。數字化就是將許多物理形態的文獻轉化為可以度量的數字,再以這些數字建立起適當的數字化模型,把它們轉化為一系列二進制代碼的電子形態。數據化則是將電子形態的文獻數據結構化,按照一定數據格式構建成適用於可製表分析的量化形式,數據分析才是其核心。文本挖掘也稱之為“文本數據挖掘”或“文本知識發現”,通俗地講就是從海量的非結構化的文本數據中發現隱含的、以前未知的、有效新穎的規律、模式和趨勢。隨著數據分析研究的不斷發展,數據挖掘的對象已遠遠超出了結構化數據的範圍,數據來源還可能是數據倉庫、文本數據集合、數據集、文檔等。面對眾多半結構化或非結構化的文本,如何有效地進行文本挖掘已經成為數字人文領域研究的趨勢之一。國外學者主要在文本結構分析、文本摘要、文本分類、文本聚類、文本關聯、文本規則和趨勢預測等方面進行了研究,比如作者的活動軌跡與歸屬、作品的情感分析、人物關係挖掘、人文學科本體構建等。<sup>⑦</sup>

21 世紀以來,隨著互聯網和數字技術的突飛猛進和迅速普及,中國的數據庫建設也進入群雄並起的階段。《2011 中國數據庫軟件市場研究報告》顯示,2007~2011 年,中國整體數據庫市場規模穩中增長,國產數據庫軟件市場每年增長的速度都在 15% 以上,呈現爆發式增長態勢。為了擺

脫約翰·內斯伯特稱之為“信息豐富而知識貧乏”的窘境,文本挖掘技術應運被引介到了國內。文本挖掘理念引導了數據庫建設、開發思路的轉變,這使得“專題數據庫”的發展呈現出與以往完全不同的面貌。例如台北“中研院”近代史所所建“胡適檔案資料庫”、“袁氏家藏近代名人手書”、“徐永昌日記”,北京中國社科院近代史所所建“顧維鈞檔案數據庫”,北京大學圖書館所建“陳翰笙檔案數據庫”,上海圖書館所建“盛宣懷檔案知識庫”,浙江大學蔣介石與近代中國研究中心籌建的“蔣介石資料數據庫”等,可謂百花齊放,異彩紛呈。更為重要的是,“人文學研究者不再是被動選擇既有的數據庫,而是參與數據庫建設過程,由其自身研究需要引導數據庫開發,數據庫開發過程也就成為其研究的一部分”。<sup>48</sup>台灣大學數字人文研究中心開發的“台灣歷史數位圖書館(THDL)”,主要收錄台灣地區契約文書及台灣總督府檔案。THDL不僅提供詞頻分析、關聯分析和人物關係分析,而且為用戶提供研究環境並幫助發現問題。香港科技大學李中清—康文林團隊建構的“財富信息數據庫”,包括中國多世代人口數據庫—雙城部分、中國土地改革數據庫—雙城部分、中國四清階級成分數據庫,“這三個子數據庫涵蓋了從19世紀到20世紀中期150年內,國家在不同的區域、社群內對財富分配進行主導、干預,以建構不平等與平等的歷史進程,為從基層出發理解中國社會的財產權利與政治權利提供了可能”。<sup>49</sup>王業鍵不僅主持建設了“清代糧價資料庫”,而且通過對清乾隆時期糧價的統計分析,發現清代官書中言之鑿鑿的“乾隆十三年米貴問題”很難成立。陳志武、彭凱翔等通過建立“中國利率史數據庫”發現,傳統借貸市場有多個層次,簡單地擷取某些樣本或者資料去評判整個借貸市場是不可取的。<sup>50</sup>文本挖掘已經切入現代文學研究:包括現代文學版本圖像史料在內的諸多文獻史料研究、作家生平傳記研究、文學社團與思潮流派的譜系研究、文學思想史與觀念史的研究、文學文本的風格學和情感分析、現代文學的跨學科與跨語言研究等。<sup>51</sup>文本挖掘進入史學領域發揮了三重功能:第一是“驗證”,借由計算與分析,量化地驗證過去史學研究者曾提出過的研究結論;第二是“修正”,透過巨量數據的分析挖掘,對過去史學研究結論進行補充修正;第三是“創新”,通過巨觀且複雜的計算,提出新的研究問題。<sup>52</sup>

#### (四) 從文本到圖形圖像可視化的飛躍

近年來,把用文字符號表示的知識轉化為圖形、圖像呈現的知識已成常態,這預示著我們已經進入讀屏、讀圖時代。可視化是利用計算機圖形學和影像處理技術,將數據轉換成圖形或者圖像在屏幕上顯示,並進行交互處理的理論、方法和技術。可視化的本質在於針對海量文本信息,最大程度地實現抽象和概括。它不僅僅是多樣的圖形、圖像的組合,更大的作用在於發現一篇文檔或者一系列文檔集合中特定的、潛在的模式結構。大體上分為三類:一是以詞頻統計方法為基礎,發現特定的詞頻模式來實現文本的可視化。二是基於聚類演算法比較文檔的相似性來實現文本的可視化。三是基於語義的標注、計算、統計等方法,發現特定的隱含語義關係來實現文本的可視化。<sup>53</sup>

數字人文的出現促進了傳統人文科學與新技術的融合,為大規模古籍可視化呈現與挖掘研究提供了可能。(1)數據可視化。以詞語首見年代為視角,以可視化(散點圖)的形式展現字詞的年代及詞頻分佈,力求量化地呈現不同時期的詞彙分佈趨勢與詞義演變規律。(2)空間信息可視化。以古籍作者為主線,利用GIS技術,對我國龐大且分散的數字化古籍進行大數據的集成和地圖展示,並呈現相關作者的地理分佈與活動軌跡,以深度開發中國古籍中的多元文化價值。<sup>54</sup>中南民族大學王兆鵬教授主持的《唐宋文學編年地圖》,經過近10年的努力,已收入唐宋時期大多數知名作家及作品的相關史料,並根據最新的研究成果及時予以更新。通過時間、人物、地點、事件四個部分,立體、多元地創建了一份時間與空間上相互銜接的文學地圖。這個項目首先帶來文學呈現方式

的三個變化：一是變選擇性呈現為全景性呈現；二是變間斷性呈現為連續性呈現；三是變未知為可知。其次是帶來文學史認知方式的變化，即從限知視角到全知視角的變化。王兆鵬認為，“可視化將是未來文學研究的必然趨勢，會成為文學研究中一個新的學術增長點”。<sup>55</sup>浙江大學徐永明教授利用 ArcGIS、QGIS、CHGIS、CartoDB、Worldmap 等數據庫和地理信息系統軟件，以湯顯祖為例，展示了其行跡和活動地點、社會關係地理分佈、社會關係點線表達等多種可視化效果圖，給人以直觀明了、耳目一新之感。<sup>56</sup>數據可視化呈現也為歷史研究提供了新視角，如歷史檔案領域可視化分析研究的“淡新檔案訴訟關係圖”、基於符號分析法的宋代政治網絡可視化研究、基於 GIS 的城市發展歷史可視化展示等，這些探索使歷史研究也可獲得直觀形象支持。數字歷史是面向公眾開放式實踐項目，旨在公眾與歷史研究之間建立互動式的溝通機制，在線互動、可視化是主要特點。<sup>57</sup>可視化方法更被應用在博物學、藝術學、文化遺產數字化保護等方面，舞蹈、音樂、語言、手稿等都是數字化的對象，目標是建設數字化的文化遺產館藏。越來越多的機構、學者借助 GIS 技術進行靜態和動態的可視化研究和呈現，如故宮博物院開發的“虛擬紫禁城”，就產生了很好的人文形象效果。

#### 四、結論：數字人文應用的價值與悖論

關於數字人文的概念，目前學術界大體上存在四種不同的代表性觀點，分別將其理解為一種實踐，看作一種研究方法，視為一種研究領域，看成一門學科。數字人文概念的提出才有十幾年的時間，但要追根溯源卻有 70 年的歷史，大體上經歷了人文計算與數字人文兩個階段。數字人文的興起不僅為人文學者提供了新的研究路徑與方法，而且為科研人員跨學科研究開辟了新的機遇和空間。“很多自然科學家開始關注人文社會科學中的問題，很多人文社會科學家更加注重不同研究取向和方式的結合”。<sup>58</sup>在數字技術與傳統人文學科研究的跨界、融合過程中，學術團體、學術會議和學術期刊構成了組合性的學術共同體，並推動了數字人文的興起和發展。數字人文應用於傳統人文學科研究的價值，主要體現在前文所述四個方面的“飛躍”。武漢大學數字人文研究中心主任王曉光教授認為，目前國際數字人文研究的前沿和典型應用包括：(1) 歷史學方面的基於 GIS 的歷史地理可視化；(2) 文學方面的文本挖掘與 TEI 標準；(3) 語言學方面的基於大型語料庫的語料庫語言學；(4) 舞蹈方面的視頻捕捉、運動分析與虛擬實境實現；(5) 考古學方面的圖像分析、色彩還原和數字重建；(6) 數字圖書館和網絡數據庫。<sup>59</sup>研究方法的革新，研究領域的拓展，研究問題的交叉，使得數字人文的內涵與外延變得更加複雜和多樣。

數字人文自誕生以來，就伴隨著不斷爭議和批評。在美國眾多反對的聲音中，有三篇文章具有代表性。2014 年，亞當·克思奇在《新共和》雜誌發表《科技接管英文系：數字人文的虛假承諾》，直接質疑高校管理的技術化和職場化，而數字人文即是其直接表現。<sup>60</sup>2016 年，丹尼爾·阿林頓在《洛杉磯書評》發表《新自由主義工具（及檔案）：數字人文的政治史》，認為數字人文並不是其所號稱的使用數字或量化方法來回答人文領域中的研究問題，只不過是將技術知識定義為一種人文知識，並認為這種知識會高於其他形式的知識，推行以計畫和 Lab 為中心的學習和研究，以取代閱讀與寫作。<sup>61</sup>2017 年，蒂莫尼·布倫南在《高等教育紀事報》發表《數字人文的幻滅》，大膽直白地批評數字人文研究對算法的依賴，認為“與其說數字人文是一場革命，不如說數字人文為了反對主流形式，從而強行將人文從其存在原因中剝離出來的那個楔子”。<sup>62</sup>這些批評提供的角度與思考，雖然另類，但頗具衝擊力。倘若作理性思考，不難發現數字人文應用於傳統人文學科研究，確實存在若干悖論。一是數字人文研究的大多數項目確實沒有推翻傳統意義上的預設，還沒有產生全新的敘

述,更沒有出現震動學界的成果。二是某些數字人文倡導者重數據輕思想、重圖像輕文字、重印證輕發現,推崇數據中心論,從而使人文研究失去人文意義,削弱甚至消解了研究者的主體性。三是GLS建設需要相對完備、充裕與精準的空間、屬性數據及規範化處理,這卻是歷史地理研究中最欠缺的。歷史地理數據的時空關係複雜,並且具有不確定性,這將使“一切皆可量化”的理想與數據獲取現實之間存在較大差距。總之,技術的每一次進步,都帶著利弊的兩重性,我們在享受數字人文帶來的各種便利時,也要避免成為“計算”和“數字”的附庸。

數字人文與傳統人文學科的聯姻已成事實,是自由戀愛還是包辦婚姻,已不重要。重要的是,一個幽靈,數字人文的幽靈,從歐美正向或正在中國遊蕩。這將預示著,新技術與古老學科的哲學關係將被重新審視,一種新的研究範式與方法不僅改變著傳統人文學科的研究手段,還改變著人文學科“解釋性學術”思考問題的角度和探索途徑。

①⑤Gerhard Lauer:《文化的數字丈量:“數字人文”下的人文學科》,龐娜娜譯,澳門:《澳門理工學報》,2018年第3期。

②⑦⑥①姜文濤:《數字人文文學研究與近代文學的興起——以北美文學研究學科史為中心的文獻學考察》,見陶東風主編:《文化研究》(第33輯),北京:社會科學文獻出版社,2018年。

③⑰凱薩琳·菲茨派翠克:《人文學科,數字化》,朱豔譯,見陶東風主編:《文化研究》(第17輯),北京:社會科學文獻出版社,2013年。

④潘威:《“數字人文”背景下歷史地理信息化的應對——走進歷史地理信息化2.0時代》,昆明:《雲南大學學報》,2018年第6期。

⑥戴安德、姜文濤:《數字人文作為一種方法:西方研究現狀及展望》,趙薇譯,濟南:《山東社會科學》,2016年第11期。

⑦凱特·巴雷特:《從小冊子到像素:轉變中的人文學科》,轉引自《西方人文社會科學研究困境與出路》,北京:《中國社會科學報》,2012年6月5日。

⑧林穗芳:《羅伯托·布薩和世界最早用計算機輔助編輯的巨著〈托馬斯著作索引〉》,河南開封:《河南大學學報》,2007年第4期。

⑨王萍、宋雪雁:《EAD、DC、TEI 著錄實例及其比較分析》,北京:《圖書情報工作》,2006年第12期。

⑩⑳蘇珊·霍基:《人文計算的歷史》,葛建鋼譯,見陶東風主編:《文化研究》(第17輯)。

⑪孟建、胡學峰:《數字人文:媒介驅動的學術生產方式變革》,北京:《現代傳播》,2019年第4期。

⑫趙生輝:《國外檔案領域數字人文項目的實踐與啟示》,杭州:《浙江檔案》,2015年第9期。

⑬⑱李子林、王玉珏等:《檔案與數字人文的“和”與“合”——國外開展檔案數字人文項目的實踐》,北京:《中國檔案報》,2018年10月15日。

⑭陳寅恪:《陳垣〈敦煌劫餘錄〉序》,《金明館叢稿二編》,北京:三聯書店,2001年,第266頁。

⑮陳靜:《歷史與爭論——英美“數字人文”發展綜述》,見陶東風主編:《文化研究》(第17輯)。

⑯徐志璋:《美國國家人文基金會支持數字人文的實踐和啟示》,天津:《圖書館工作與研究》,2018年第3期。

⑰劉煒、葉鷹:《數字人文的技術體系與理論結構探討》,北京:《中國圖書館學報》,2017年第5期。

⑳趙生輝、朱學芳:《我國高校數字人文中心建設初探》,北京:《圖書情報工作》,2014年第6期。

㉑魏清華:《歐洲數字人文教育現狀及其啟示》,瀋陽:《圖書館學刊》,2018年第5期。

㉒⑳郭晶、王曉陽:《國外數字人文研究演進及發展動向——基於哈佛大學圖書館館藏相關專著的梳理》,蘭州:《圖書與情報》,2018年第3期。

㉓托馬斯·庫恩:《必要的張力》,范岱年等譯,北京:北京大學出版社,2004年,第288~289頁。

㉔中國科協主席韓啟德致第十一屆中國科協年會開幕詞,人民網,2009年9月8日。<http://scitech.people.com.cn/GB/25509/56813/166862/166864/10009618.html>

㉕李劍鳴:《自律的學術共同體與合理的學術評價》,北京:《清華大學學報》,2014年第4期。

㉖㉗周晨:《國際數字人文研究特徵與知識結構》,廣

- 州:《圖書館論壇》,2017年第4期。
- ⑳ Susan Schreibman, Ray Siemens, John Unsworth, (ed.) *A companion to digital humanities*, Oxford: Blackwell, 2004, p.3-19.
- ㉑ 高瑾:《數字人文學科結構研究的回顧與探索》,廣州:《圖書館論壇》,2017年第1期。
- ㉒ 朱劍:《學術共同體、學術期刊體制與學術傳播秩序——以媒體更迭時代人文社會科學期刊轉型為中心》,澳門:《澳門理工學報》,2016年第3期。
- ㉓ 張斌:《我國學術共同體運行的現狀、問題與變革路徑》,北京:《中國高教研究》,2012年第11期。
- ㉔ 原祖傑:《學術期刊何以引領學術——兼論學術期刊與學術共同體之關係》,澳門:《澳門理工學報》,2014年第1期。
- ㉕ 王濤:《中國的世界史研究實現數字轉型要面對的問題》,南京:《南京大學學報》,2019年第2期。
- ㉖ 劉煒、林海青等:《數字人文研究的圖書館學方法:書目控制與文獻循證》,北京:《大學圖書館學報》,2018年第5期。
- ㉗ 馮惠玲:《數字人文——在跨界中實現交融》,北京:《中國社會科學報》,2017年12月21日
- ㉘ 申斌、楊培娜:《數字技術與史學觀念——中國歷史數據庫與史學理念方法關係探析》,北京:《史學理論研究》,2017年第2期。
- ㉙ 梁晨:《量化數據庫與歷史研究》,北京:《歷史研究》,2015年第2期。
- ㉚ 雷家瓊:《構建量化數據庫 算出更精確的歷史——訪香港科技大學人文社會科學學院院長李中清教授》,北京:《中國社會科學報》,2014年6月30日。
- ㉛④④ 包弼德:《數字人文與中國研究的網絡基礎設施建設》,夏翠娟翻譯整理,上海:《圖書館雜誌》,2018年第11期。
- ㉜ 梁晨:《量化數據庫:“數字人文”推動歷史研究之關鍵》,南京:《江海學刊》,2017年第2期。
- ㉝ 王大學:《國際學界國家歷史地理信息系統建設與利用的現狀及啟示》,徐州:《江蘇師範大學學報》,2016年第3期。
- ㉞ 廖泮銘、范毅軍:《中華文明時空基礎架構:歷史學與信息化結合的設計理念及技術應用》,北京:《科研信息化技術與應用》,2012年第4期。
- ㉟ 張萍:《地理信息系統(GIS)與中國歷史研究》,北京:《史學理論研究》,2018年第2期。
- ㊱ 郭金龍、許鑫:《數字人文中的文本挖掘研究》,北京:《大學圖書館學報》,2012年第3期。
- ㊲ 趙思淵:《地方歷史文獻的數字化、數據化與文本挖掘:以〈中國地方歷史文獻數據庫〉為例》,北京:《清史研究》,2016年第4期。
- ㊳ 梁晨:《用大規模量化歷史數據庫檢驗中國的長期代際遺傳》,南京:《南京大學學報》,2019年第2期。
- ㊴ 陳志武、彭凱翔:《清初至二十世紀前期中國利率史初探——基於中國利率史數據庫(1660~2000)的考察》,北京:《清史研究》,2016年第4期。
- ㊵ 王賀:《“數字人文”如何與現代文學研究結合》,上海:《現代中文學刊》,2019年第1期。
- ㊶ 邱偉雲:《驗證、修正、創新:數字史學方法的三重功能》,南京:《南京大學學報》,2019年第2期。
- ㊷ 張智雄、孫坦:《文本可視化及其主要技術方法研究》,北京:《現代圖書情報技術》,2008年第8期。
- ㊸ 歐陽劍:《面向數字人文研究的大規模古籍文本可視化分析與挖掘》,北京:《中國圖書館學報》,2016年第2期。
- ㊹ 王兆鵬:《今後古典文學研究的可視化趨勢》,北京:《古代文學前沿與評論》,2018年第1期。
- ㊺ 徐永明:《中國古典文學研究的幾種可視化途徑——以湯顯祖研究為例》,杭州:《浙江大學學報》,2018年第2期。
- ㊻ 張耀銘:《人工智能驅動的人文社會科學研究轉型》,濟南:《濟南大學學報》,2019年第4期。
- ㊼ 王曉光:《“數字人文”學科的產生、發展與前沿》,見全國高校社會科學科研管理研究會組編:《方法創新與哲學社會科學發展》,武漢:武漢大學出版社,2010年。
- ㊽ 李點:《面對數字人文的幽靈》,河北秦皇島:《燕山大學學報》,2017年第1期。
- ㊾ 轉引自陳靜:《數字人文知識生產轉型過程中的困境與突圍》,見陶東風主編:《文化研究》(第33輯)。

作者簡介:張耀銘,《新華文摘》雜誌社編審。  
北京 100706

[責任編輯 劉澤生]