

文化的數字丈量： “數字人文”下的人文學科

〔德國〕Gerhard Lauer（撰）龐娜娜（譯）

〔提 要〕 作為一種新興的人文學科的研究方法，“數字人文”起源於20世紀末，這一方法基於對電腦和互聯網的運用，是一種更廣闊意義上的“計算”。“數字人文”的概念起源於語文學的傳統，是人文學科的延展。“數字人文”的發展依託於數據和元數據的標準化，其研究建立在大量的數據基礎、統計學計算、可視化圖表之上。“數字人文”使人文學家以一種新的方式丈量文化，大大擴展了人文學科研究的廣度和深度，但也面臨著諸多問題與挑戰。

〔關鍵詞〕 數字人文 語文學 人文學 人文計算

〔中圖分類號〕 I0-05;G05 **〔文獻標識碼〕** A **〔文章編號〕** 0874 - 1824 (2018) 03 - 0132 - 08

一、引言

如果說變革的意義在於改變生活，那麼數字化就是這樣一場變革，它急劇地改變了我們生活的方方面面，而且風頭正勁。數字化也改變了人文學科，^①這場變革以“數字人文”為標籤，但“數字人文”這一概念正如同許多學科名稱一樣尚不精確。1960年代，電腦被運用於人文學科研究中，人們將其暫稱作“人文計算”。直到世紀之交，隨著互聯網的廣泛傳播與應用，人們才把這種依靠電腦進行研究的人文學科稱作“數字人文”。這一領域出版物的名稱也隨之相應發生變化：早在1980年代，該領域出版物尚稱為《人文計算年鑑》(*The Humanities Computing Yearbook*)，到了2000年已更名為《數字人文手冊》(*A Companion to Digital Humanities*)。隨著名字的變更，這一領域涵蓋的學科也更加廣泛，它不僅涵蓋純粹的文字學科，也包括歷史學、考古學、音樂學和藝術學等學科。在歐洲，這一領域相關學術協會的名字也發生了變化，由“文學和語言學計算協會”(Association for Literary and Linguistic Computing)更名為“歐洲數字人文協會”(The European Association for Digital Humanities)。這一領域內的雜誌也與協會同步，紛紛更名。數字人文專家威廉·麥卡蒂(Willard McCarty)^②與馬修·柯申博姆(Matthew Kirschenbaum)^③認為，“數字人文”已經超越了人文學科原本的研究範疇，並且在人文學科中變得越來越顯性。^④

名字的變更並不僅僅意味著人文學科的下一個“轉向”，人們更關注的問題還是人文學科的價值等級（Wertehierarchie），即人文學科一直探討的什麼在人文研究中更重要的問題。長期以來，人文學科普遍認為，研究對象比研究方法更為重要。因此，“計算”一直以來都不屬於人文學科的研究方法，並且研究方法的創新在人文學科一直未得到應有的重視。在自然科學和生命科學領域，人們對電腦的使用頗為普遍，對一些以電腦為依託的學科如計算物理學、生物信息學、地理信息學等早已司空見慣。1986年諾貝爾物理學獎將獎項授予格爾德·賓寧（Gerd Binnig）和海因里希·羅勒（Heinrich Rohrer），以表彰他們對掃描隧道顯微鏡（scanning tunneling microscope）^⑤做出的發展。這也體現了這些學科對研究方法創新的高度重視。

然而，在人文學科，這一情況卻截然相反。人文學科歷史、闡釋性的研究路徑源自古希臘的文本批評與14至16世紀的人文主義。它可以追溯到以克里斯蒂安·戈特洛布·埃內（Christian Gottob Heyne, 1729~1812）和弗里德里希·奧古斯特·沃爾夫（Friedrich Augustus Wolf, 1759~1824）為代表的古語言學及19世紀以語言學家威廉·狄爾泰（Wilhelm Dilthey, 1833~1911）為代表的語言學。一直以來，這一傳統統治著人文學科。雖然有時人文學者也有對現代化和創新的願望，但這種願望的實現也僅停留在對人文學科表層轉向的滿足。他們認為，文化評判是溫暖的、充滿人文關懷的，與冷酷的科技文明截然不同。

基於計算的人文學科就是在這樣的背景下作為一種矛盾體出現的。“計算”這種方法並不能立即被人文學科作為研究方法採納，因為這一研究路徑超越了定性闡釋的研究範疇，因此數字人文是一個便利的標籤。這一概念的不確定性與廣闊性解決了這一學科在研究方法和學科認識上的適應困難。這就是“數字人文”：它是人文學科研究方法的補充，這一方法基於對電腦和互聯網的運用，是一種更廣闊意義上的“計算”。^⑥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”。“數字人文”並非屬於計算機學科的轉向，它更多的屬於人文學科。數字人文的論壇“人文主義”、“德語區的數字人文”以及這一學科的專家達恩·科昂（Dan Cohen）^⑦和馬丁·米勒（Martin Mueller）^⑧都認為，“數字人文”是人文學科一種謙虛的說法，與語文學一直以來謙遜的傳統有關。^⑨

二、數字人文與語文學的淵源

語文學開始於闡釋前，也就是收集數據（數據可能是文本或者是待研究物），然後將數據區分、分類、對比、關聯、整理。這種基本的研究方法構成了人文學科的基礎。^⑩這種方法的運用體現在特奧多爾·蒙森（Theoder Mommsen, 1817~1903）對全部羅馬帝國銘文的搜集上，他按照省份對這些銘文進行了分類整理。1853年，在蒙森開始這一里程碑式的項目時，他並沒有提出具體的研究問題，只是想讓這一數據可以供後來的研究者使用。如今，拉丁銘文集數據庫已經涵蓋超過18萬則銘文。這一數據庫的建立標誌著科學史向新時期的轉變：這是一種新形式的處理數據的方式。歷史學家洛蘭·達斯頓（Lorraine Daston）^⑪展示了如何在現代化的技術支持下對異國珍品進行可調控精度的觀察，這些藝術品原本只會出現在藝術典藏室或拍賣場上。正是這些創舉而非闡釋學，讓科學也讓語文學變得現代化。數字人文將繼續使用這些路徑，並對其做出調整。

蒙森並不是唯一一個運用這種方法收集數據的學者。差不多同時，英國數學家奧古斯都·德·摩根（Augustus de Morgan, 1806~1871）產生了一種思路：通過計算音節數量來區分作家寫作風格。摩根想要考察聖經中的一些信件是不是真的由使徒保羅（Paulus, 約3~67）所撰寫。為了判

斷這些信件的作者，他沒有運用傳統闡釋學的路徑，而是確定每封信中詞的平均長度。他的預設很簡單：一位作者在信中運用的詞長應該與別的作者運用的詞長有所區別。

幾年之後，美國氣象學家兼狂熱的莎士比亞迷托馬斯·科溫·門登霍爾（Thomas Corwin Mendenhall, 1841~1924）嘗試描繪英國作家的寫作特徵。門登霍爾並不考察文本中的詞長，而是研究均勻劃分的段落中詞匯出現的頻率。他的這一研究理清了莎士比亞與他同時代的作家培根之間寫作風格的區別。後來，門登霍爾又用這一研究方法對狄更斯、薩克雷與密爾的作品的文體特徵進行了研究。

與摩根和門登霍爾的研究不同，19世紀末，波蘭古語文學家、哲學家溫切蒂·盧托斯勞斯基（Wincenty Lautoslawski, 1863~1954）嘗試將有爭議的柏拉圖的作品年表進行界定，他通過計算統計不同形式的文本特性來達到這一研究目的。盧托斯勞斯基從這種假設出發，即作家作品形式的特徵是逐漸發展而來的，因此在大致時間內產生的文章應該會在形式特徵上顯示出極大的相似性。盧托斯勞斯基甚至認為，人們可以判定文體計量的模式（Stilometrische Muster），並通過這種模式區分作家，而不是通過傳統闡釋學的路徑。^⑫

這些考察作家文體特徵的相似的研究，不僅出現在19世紀語文學的傳統中，也出現在俄國的文學研究和法國的結構主義中。今天，這一研究方法依舊發揮著重要的作用。比如，在判定作品作者的研究中，在法醫、刑偵方面的運用，在注重調查的新聞業及國家安全方面的應用。^⑬

雖然這些語文學領域的研究與數學關係緊密，但其中的一些研究已經可以歸為語文學的核心事件——文本的版本問題。1949年，當僅有少數的文學研究者知道有一種東西叫電腦時，神父羅伯托·布薩（Roberto Busa, 1913~2011）在他的文章中闡釋聖托馬斯“存在”這一概念時就談到：電腦也許是一個有用的工具，可以用來瞭解這個幾乎沒法瞭解的大塊頭作品 *Thomisticum* 中轉換概念的運用，比如確定兩個詞共同出現的情況、展示他們平行出現的位置（Parallelstellen）等。布薩的研究問題和研究方法並不新穎，但是在這一研究中他對電腦的運用卻是革命性的。布薩神父成功地讓IBM的創始人托馬斯·沃森（Thomas J. Watson）對他開創性的研究理念感到驚奇。在IBM的合力幫助下產生了第一個穿孔卡片上的電子版本的 *Thomisticum* 數據庫。它以一種全新的方式解析作品，是一個文化歷史意義重大的工程，如今已經有56卷紙版和電子版供讀者參閱。1960年代出現了第一個電子版本的“現代語言協會國際書目”（Modern Language Association International Bibliography），它囊括了現代語言學各個階段的專業書目。研究者可以通過用於數據傳輸的電話網絡進行檢索。這個數據庫中包含了很多作品，如格林兄弟的《德語大詞典》、歷史詞典 Krünitz 及一些地區性方言詞典。誰想查閱某個概念的使用，無論這個概念是“靈魂”（Seele）、“大腦”（Gehirn）“土耳其”（Türkei）還是“宙斯”（Zeus）、“啤酒乳清”（Biermolke）、“透翅蛾科”（Glasflügler），都可以登錄 Woerterbuchnetz.de 這個網站，在上面可以查到這個詞的近義詞、反義詞、詞源等信息。雖然這是人文學科幾百年來一直使用的基本操作方法，但電子詞典和電子目錄正使人文學科向不可預測的廣度和深度發展。數碼世界擴展和加深了“博學”（Gelehrsamkeit）這一概念的內涵。

許多關於文學作品的數據庫隨後逐漸產生，從古斯塔夫·福拜樓的小說到沃爾特·惠特曼的詩歌、女性文學、16世紀的印刷品、愛爾蘭家族譜系表，甚至到中世紀流傳下來的沒有印刷版的 *Parzival*^⑭手稿。這些項目構成了近30年數字人文的研究的主要內容。當然這些研究中的很多可以與文化經典的電子化與電子編輯掛鉤。古語言學家格雷戈里·克蘭（Gregory Crane）建立了玻

耳修斯數碼圖書館供研究使用，這是一個集古希臘、古羅馬藝術品的電子庫；1999年由卡爾·埃布（Karl Eibl）和福提斯·雅內蒂斯（Fotis Jannidis）出版的第一個數字人文專業雜誌以“計算機語文學年鑒”為標題；谷歌開始對書籍進行大規模的數字化轉換。歷史上還從來沒有這麼多關涉語文學的研究，這要歸功於數字人文這一學科的持久影響力。

數字化版本一直是這一領域的核心議題。近年來也出現了越來越多數字化的文化實體。玻耳修斯圖書館展覽著數千件如古幣、花瓶、雕塑等珍貴的古代藝術珍品。圖畫和音樂作品也被列為了數字化的對象，比較著名的有和惠普公司合作的莫紮特數字音樂庫（Mozart-Ausgabe），還有藝術史學家馬丁·肯普（Martin Kemp）的“廣博的達芬奇”（Universal Leonardo）項目。達芬奇項目中包含了達芬奇藝術品和科技作品的電子版。這些作品的電子版不僅能夠提供任何一個實體博物館都不能提供的高解析度，而且可以通過X射線、紅外線等技術使研究者的觀察更加深入化。^⑩這些電子版對研究的益處是顯而易見的：莫紮特電子音樂庫每天被全世界的愛好者使用。藝術品的鑒賞基礎發生了改變，因為現在任何一個高年級的學生都可以在數據庫中流覽文藝復興時期的珍貴手稿；任何一個本科生都可以在寫論文時使用這些原來對他們遙不可及的珍貴圖片。文化這一概念正在悄然發生變化，因特網上的文化產品正在覺醒。^⑪

三、數字人文的研究範疇

數字人文和數碼編輯能夠成功的一個重要前提就是數據和元數據的標準化。在文本編碼計畫（Text Encoding Initiative）框架下，人們費時幾十年研發了一組標準數碼標記辭彙表，這套標準對不同語言中的不同文本類型和文獻形式統一適用，它使文本不再受制於各個軟、硬件的不同賦碼標準。其他學科也緊隨這一趨勢，音樂學科研發了統一賦碼標準MEI（Music Encoding Initiative）。圖書館、數據中心、萬維網聯盟（W3C）為了能夠將這些電子版本納入他們的目錄與伺服器對這些元數據尤其重視。數據標準化和多個機構的協同合作是數字人文學科的典型特徵。甚至可以這樣說：數字人文必然是國際化的。

雖然取得了不俗的成果，然而，近年來，數字人文因研究過於保守而頗遭詬病。數字人文研究總是以單個文本或者文本電子化為導向，並未脫離歷史闡釋學的傳統，^⑫美國文學批評家約翰·克勞·蘭瑟姆（John Crowe Ransom）把這一模式稱為“細讀”（close reading）。細讀指的是將作品中的辭彙組合、詞義等逐句、逐段詳細分析。人文學者總是固守於經典作品中故事展開的特殊藝術模式，對關涉文本意義的上下文、語用信息甚至文化、歷史信息都拋諸腦後。

也許沒有人比美國比較文學家弗蘭克·莫雷蒂（Franco Moretti）對這一傳統的研究路徑批評得更為中肯。2000年，他在《新左派評論》（New Left Review）雜誌上發表了名為《對世界文學的猜想》（“Conjectures on world literature”）的文章，引發了廣泛的討論。在該文中，莫雷蒂關注的不是解決數字人文的概念問題，而是尋找他所在領域的研究方法。沒有人能夠讀完所有作品，也不可能有人理解世界上的所有語言，那麼在這種情況下，人們又該如何撰寫世界文學史呢？作為對這一問題的回答，他提出了“遠讀”（distant reading）的概念。他在文章中寫道，文學史將很快脫離現在的樣子，它將變成“二手”的文學史，將成為由不同學者的研究成果構成的“拼貼品”（Patchwork），更確切地說文學史中不再會出現單一文學作品的文本閱讀。莫雷蒂的這一說法要比“世界文學”（Weltliteratur）的概念走的更遠，他將研究項目的雄心與文本的距離直接相連：研究項目的雄心越大，那麼與文本的距離就越遠。^⑬

莫雷蒂的這一提法與傳統文學研究徹底分道揚鑣，他的同事斯坦利·費什（Stanley Fish）對他這個雄心勃勃的“遠讀”的計畫頗有微詞。^⑩但莫雷蒂的這一提法卻迅速傳播開來，因為它很好地概括了數字時代我們面臨的挑戰：我們現在面對著成百萬、上千萬的圖書。數字人文原來迫不得已只能研究單個經典作品，因為那時候大規模的歷史性作品尚未被存儲。甚至到20世紀末，《格林大詞典》還受制於CD的限制。但是隨著因特網的發展、存儲空間的降價以及電子化工具的增加，人文學科中的大數據已經不僅僅只是一種隱喻了。根據谷歌推斷，全世界大約有1.3億冊已經出版的圖書。谷歌已經成功將其中的2,000萬冊數據化。當谷歌2004年啟動這一計畫時，沒有人會想到，僅僅是過了十多年的時間，人們現在用鼠標輕輕一點，就能檢索到如此數量的書籍。在其他的數據庫中，如在線的“早期英語圖書”（Early English Books）、Evans、Gallica、德國文本檔案（Deutsches Textarchiv）、TextGrid Repotorium等數據庫^⑪中都有數百萬計的圖片、印刷品和書籍。誰想要找18世紀人類學第二版的一個小冊子、或者是一份波斯手稿、手寫成績單、16世紀米歇爾·德·蒙田（Michel de Montaigne, 1533~1592）的肖像，幾秒鐘就能在這些數據庫中找到。科學家、政治家和民眾將日本海嘯災難留下的所有媒體痕跡——如圖片、推特、新聞報導等匯總到“2011年日本災難電子檔案”中，通過這種方式建立了一個流動性的世界性事件的電子檔案。

數字人文學家以一種新的方式丈量文化。考古學給出了很好的例子：原來科學家發現一個雕塑的某個部分，會先用石膏建模，然後帶著這個石膏模型去世界各個博物館考察，看這個頭或者那個胳膊屬於哪一尊雕像。如今這一工作通過3D掃描機就可以完成。之前必須周遊各大洲才能完成的工作，如今運用因特網就可以完成。鐳射掃描器可以將地形掃描並以一種從未有過的精度製圖。達米安·埃文斯（Damian Evans）和讓-巴蒂斯特·塞爾文斯（Jean-Baptiste Chevance）用這種方法在柬埔寨吳哥窟以北40公里處發現了千年古城瑪漢德拉帕瓦塔（Mahendraparvata）。那個地方肉眼看上去就是一片熱帶叢林，但是電腦顯示了地下岩石層和儲水情況。歷史學家也開始運用數字手段進行研究。比如歐洲大屠殺研究中心（EHRI）致力於用波蘭收稅單、德國人驅逐名單、影像資料等證據來還原在二戰期間被屠殺的600萬猶太人的姓名。目前，通過這種方法，他們已經還原了三分之一的姓名。不同來源的數據會被處理，然後以可檢索的方式存儲在數據庫中。這些數據的處理和存儲都需要運用電腦。上述這些數字人文項目不僅僅是人文學科的延展，它也開啟了一個新的研究世界。

四、數字人文的發展與面臨的挑戰

人文學科當下的轉型較少涉及文化研究借助新的研究方法所能達到的深度，而是更多涉及其廣度。以前文化的存儲量是有限的，人們只能局限於對經典作品的研究。如今，數字時代幾乎使一切文化都觸手可及。如果我們對數字人文接下來的發展方向做出預測的話，我們認為：這將是一場文化革命。

借助數字方法，我們現在突然能夠理解制陶技術能夠在不同時代、不同的文化中廣泛傳播的原因。我們可以發現，它的傳播是否與傳染病的傳播路數相近、是否沿著貿易道路進行。荷蘭文化史學家喬普·萊爾森（Joep Leersen）想要瞭解，19世紀民族國家這一概念是以何種方式傳播的，有多少人曾參與其中。他採用研究幾個人物的書信往來及信件交匯點的方法，並把研究結果用圖表可視化，發現歐洲北端的愛丁堡與南端的薩拉熱窩都被涵蓋進網絡中。文學批評家約翰·伯羅

斯（John Burrows）、朱莉婭·佛蘭德斯（Julia Flanders）、大衛·胡佛（David Hoover）、福提斯·雅內蒂斯（Fotis Jannidis）、馬修·約克斯（Matthew Jockers）^①等致力於研究如下問題：時代印象僅僅是由歷史圖像構建的嗎？愛爾蘭移民文學何時開始向美國遷移的？19世紀出現了哪些新的主題？性別對於作家的寫作風格究竟有哪些影響？克萊斯特（Heinrich von Kleist, 1777~1811）與他同時代的古典主義和浪漫主義時期的作家的寫作風格完全不同嗎？對這些問題的回答建立在大量的數據基礎、統計學的計算、可視化的圖表之上。或者我們可以把這一過程稱之為“事實挖掘”（Reality-Mining）。所有這一切都大大擴展了人文學科研究的可能性。

然而，數字人文學者在進行研究時是以歷史性、闡釋性的知識為前提的，他們需要理論模型，因為單純的數據並不能作為研究思路和技術來運用。數據尚不是信息，信息需要有結構，而這一結構又必須從數據中推導而來。因此，“數字人文”這一學科目前急需一種能夠將文化各方面的研究和語義實體方面的各項成果都涵蓋進來的理論。歐洲數字圖書館（Europeana）的歐洲數據模型及萬維網聯盟（W3C）的資源描述框架（RDF）都是建立在語義網絡標準上的，無以計數的文化和理論知識都存在於這一標準中。如果人們想將這些數據化的文化體聯繫起來，那麼就要考慮他們出現的語境：比如，這個文件來源於這個文化空間、並且與讀者群相關；這幅畫與中國的義和團運動相關；阿弗·貝恩（Aphra Behn, 1640~1689）是第一位英國職業作家；殖民主義和近東的現代衝突之間存在著關聯等等——我們永遠也無法羅列盡這其中涉及的理論與知識。數據處理技術變得日益重要，微軟研發部的吉姆·格雷（Jim Grey）和托尼·海恩（Tony Hey）已經提到了與數據處理相關的第四個範式——數據密集型科學。這一範式在經驗、理論和模擬之後，^②該領域的方法論與理論模型變得越來越重要，數據越多，需要的理論、知識和方法論就越多。

對人文學科的數據現代化而言，方法論與理論的批判總是一個演算法的問題，是一個計算的路徑。^③新的研究方法不是由電腦來發展的，它也不會從材料中自己產生。我們該如何測量文本之間的距離？是用約翰·布羅斯（John F. Burrows）的德爾塔理論（Delta）^④還是歐幾里得演算法？生物學中哪些種系測量法最適合檢測族系的歷史文化趨勢？在描述文化親緣關係時，SplitsTree^⑤和Dendrogramm^⑥兩種表格各有何優點？所有這些方法論和理論上的問題都是人文學科面臨的問題，人文學科需要理論。

對數學演算法的探討僅僅只是數據現代化中眾多問題的一個。這還與另一情況相關，即數字人文學者能夠接觸並掌握數據，只有這樣數字人文才能成為可能。誰想研究達爾文主義的前史，將艾瑪·達爾文^⑦（Emma Darwin, 1808~1896）的日記與約翰·弗里德里希·布盧門巴赫^⑧（Johann Friedrich Blumenbach, 1752~1840）書中的豪豬刺相聯系，誰想研究早期猶太教堂的建造技術並想探究油畫配色在歐洲藝術史中的改變，就必須先接觸到這些數據。^⑨我們需要的是文化，而不是現在占支配地位的商業化和壟斷主義。谷歌圖書就是一個例子。雖然，該公司掌握著大量的數據化、機器可讀的圖書，但是這個圖書館卻存在各種不同的秘密協定。如果人們要使用博物館的藏品，就要交入館費；如果人們要研究出版社的書籍或文章，也要向圖書館繳納高昂費用，而且出版社禁止圖書館提供這些文章或書籍的電子版；國家也尋求對其文化遺產的數據高解讀性與掌控性。就連一些科學界人士在這方面表現的也不盡人意。^⑩雖然新知識在世界各個角落層出不窮，知識科技經歷著新的模式，^⑪但數據卻被日漸封閉起來。關鍵詞如“群智科學”（Crowd Science）或者“公民科學”（Citizen Science）^⑫僅僅只是美好的願景，各個“數據小聯邦”讓文化數據受到各種限制。文化和科技政策將來也會涉及到數據上，數字人文必須要與一些如“開放獲取”（Open Access）

和“知識共享”（Creative Commons）的行動相關聯。^⑨理論上來講，中世紀的日本小說距離我們很近，只需點擊鼠標，機器人就可以把它翻譯成各種文字；數學演算法也可以解密很多文化相關性，然而這些數據卻不是每個研究者都可及的。數字社會的封建主義結構與數字人文開放的學科特性背道而馳。然而，對文化的數字丈量已經開始，可以說，我們正在路上。

-
- ①人文學科（Geisteswissenschaften），在德語中，這一概念大致包括四十個單獨學科，這些學科涉及文化、思想、媒體、社會、歷史、政治、宗教等科目。——譯注
- ②威廉·麥卡蒂是英國倫敦國王學院數字人文系教授，《人文計算》一書由他編纂，此外他還在數字人文領域發表了許多學術文章。——譯注
- ③馬修·柯申博姆是馬里蘭大學副教授、馬里蘭人文技術研究所（MITH）副主任。——譯注
- ④Matthew Kirschenbaum, “What is digital humanities and what's it doing in English Departments?”, in: ADE Bulletin 150, 2010, S. 1-7, online verfügbar unter: http://mkirschenbaum.files.wordpress.com/2011/01/kirschenbaum_ade150.pdf.
- ⑤掃描隧道顯微鏡作為一種掃描探針顯微術工具，可以讓科學家觀察和定位單個原子，它具有比它的同類原子力顯微鏡更加高的解析度。此外掃描隧道顯微鏡在低溫下可以利用探針尖端精確操縱原子，因此它對於納米科技既是重要的測量工具又是加工工具。——譯注
- ⑥Peter Lunenfeld, Anne Burdick, Johanna Drucker, Todd Presner, Jeffrey Schnapp, *Digital Humanities*, Cambridge (Massachusetts): MIT Press, 2012, online verfügbar unter: http://mitpress.mit.edu/sites/default/files/titles/content/9780262018470_Open_Access_Edition.pdf.
- ⑦參見達恩·科昂網頁，<http://www.dancohern.org/2009/11/18/introducing-digital-humanities-now>。——譯注
- ⑧參見馬丁·米勒網頁，<http://literaryinformatics.northwestern.edu>。
- ⑨人文學科一直以來都認為，研究對象是重要的，而非自己的研究成果和研究方法。如某位文學評論家對歌德的作品進行闡釋，那麼歌德的作品是最重要的，評論家的闡釋僅僅是對作品的闡釋而已。——譯注
- ⑩ John Unsworth, “Scholarly primitives: what methods do humanities researchers have in common, and how might our tools reflect this?”(2000), online verfügbar unter: <http://people.lis.illinois.edu/~unsworth/Kings.5-00/primitives.html>.
- ⑪Lorraine Daston, Katharine Park (2002), *Wunder, Beweise und Tatsachen, Zur Geschichte der Rationalität*, Frankfurt am Main: Fischer, 2002.
- ⑫參見 Peter Grzvbek, (Hg.), *Contributions to the Science of Text and Language: Word Length Studies and Related Issues*, Dordrecht: Springer, 2007.
- ⑬Patrick Joula, *Authorship Attribution*, Boston/Delft: Now, 2008.
- ⑭Parzival 是德國中世紀宮廷文學的一部詩體小說，大約產生於 1200 至 1210 年間。——譯注
- ⑮Hubertus Kohle, *Digitale Bildwissenschaft*, Glückstadt: Hülbusch, 2013, online verfügbar unter: <http://archiv.ub.uni-heidelberg.de/ardok/2185/>.
- ⑯Dieter Uckelmann, Mark Harrison, Florian Michahelles, *Architecting the Internet of Things*, Heidelberg/New York/Dordrecht/London: Springer, 2011.
- ⑰Manfred Thaller, (Hg.), *Controversies Around the Digital Humanities, Historical Social Research* (37/3), 2012.
- ⑱Franco Moretti, “Conjectures on world literature”, in: *New Left Review I*, 2000, S. 54-68, online verfügbar unter: <http://newleftreview.org/II/i/franco-moretti-conjectures-on-worldliterature>.
- ⑲Matt Erlin, Lynne Tatlock (Hg.), *Distant Readings: Topologies of German Culture in the Long Nineteenth Century*, Rochester: Camden, 2013.
- ⑳詳情參見這些數據庫的網址：<https://eebo.chadwyck.com/home>, <http://childe.talkbank.org/access/Eng-NA/Evans.html>, <http://gallica.bnf.fr/accueil/?mode=desktop>, <http://www.deutschestextarchiv.de/>, <https://textgridrep.org>。——譯注

- ㉑Matthew L. Jockers, *Macroanalysis: Digital Methods and Literary History*, Urbana/Chicago/Springfield: University of Illinois Press, 2013. zusätzliches Material ist online verfügbar unter: <http://wwwmatthewjockers.net/macroanalysisbook/>.
- ㉒Tony Hey, Stewart Tansley, Kristin Tolle (Hg.), *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Seattle: Microsoft Research, 2010, online verfügbar unter: <http://research.microsoft.com/en-us/collaboration/fourthparadigm/>.
- ㉓Stephen Ramsay, *Reading Machines: Towards an Algorithmic Criticism*, Urbana/Chicago/Springfield: University of Illinois Press, 2011.
- ㉔這一理論由約翰·布羅斯提出，它可以檢測兩篇文章之間的差異，可以解決作者歸屬問題。——譯注
- ㉕SplitsTree 是一種流行的程序，用於從各種類型的數據（如序列比對，距離矩陣）推斷系統發育樹或更一般的系統發育網路。——譯注
- ㉖Dendrogram 是一種樹圖，經常用於說明通過分層聚類生成的集群排列。它通常用在計算生物學中來說明基因或樣品的聚類。——譯注
- ㉗艾瑪·達爾文是達爾文的表姐與妻子。——譯注
- ㉘約翰·弗里德里希·布盧門巴赫是一位德國醫學家、生理學家、人類學家。——譯注
- ㉙Mingquan Zhou, Guohua Geng, Zhongke Wu, *Digital Preservation Technology for Cultural Heritage*, Heidelberg/New York/Dordrecht/London: Springer, 2012.
- ㉚參見 Lauer Gerhard, "Bibliothek aus Daten", in Kodex I, S.79-85, 2011, online verfügbar unter: <http://gerhardlauer.de/index.php/download-file/view/187/85/>。
- ㉛Luciano Floridi, "Internet. which future for organized knowledge, Frankenstein or Pygmalion?", in: *International Journal of Human-Computer Studies* 43/2, 1995, S.267-274, online verfügbar unter: <http://www.philosophyofinformation.net/publications/pdf/frank.pdf>.
- ㉜Tobias Blanke, *The Ecosystem of Digital Assets, Crowds and Clouds*, Oxford: Chandos, 2014.
- ㉝Lawrence Lessig, *Freie Kultur: Wesen und Zukunft der Kreativität*, München: Open Source Press, 2006, das englische Original ist online verfügbar unter: http://www.jus.uio.no/sisu/free-culture.lawrence_lessig/portrait.letter.pdf .

作者簡介：Gerhard Lauer，瑞士巴塞爾大學數字人文中心主席，巴塞爾大學德語語言文學教授。瑞士巴塞爾 4056

譯者簡介：龐娜娜，南京大學外國語學院博士研究生，德國哥廷根大學博士研究生。南京 210023

[責任編輯 桑 海]