

“精確人文科學”的價值

[德] Gerhard Lauer(撰) 龐娜娜(譯)

[提要] 作為一種新興的人文科學研究方法,數字人文為人文科學提供了更好的研究基礎和更廣闊的發展前景。但這種“精確的人文科學”對研究方法和研究內容的革新引發了“數字人文之爭”。實際上,人文科學的精確化之路有着悠久的歷史,在人文科學研究中走更精確的道路並不是一種破壞性行為。“精確人文科學”對人文科學的進一步發展是有價值的。“精確人文科學”能否在 21 世紀的人文研究中扮演更重要的角色,取決於人文學者們對它的態度。

[關鍵詞] 數字人文 精確人文科學 不精確人文科學 價值

[中圖分類號] I0-05;G05 [文獻標識碼] A [文章編號] 0874-1824(2020)03-0144-12

一、引言

不論是在博物館還是在大學,抑或是在其它什麼地方,用魚敲打油畫都不是一個可被社會接受的行為。在接受《衛報》採訪時,蘇格蘭著名作家艾莉森·路易斯·肯尼迪(Alison Louise Kennedy)便用這樣一個“拿魚打畫”的比方,來說明計算機為人文科學和藝術領域所作的“貢獻”。在肯尼迪看來,就文學而言,計算機輔助法不能為人文和藝術領域帶來任何新東西;數字化不過是人文科學在失去前進方向的狀態下做出的錯誤嘗試。^①沒有比這更直接的批評了。隨着計算機的應用,文化研究失去了肯尼迪所謂的某種評價文學和藝術的能力。如果人們認同肯尼迪的觀點,那麼那些追求數字化的文化分析專家們的評價標準就會變得很怪異:專家們好像只字不提文學,僅僅把文學故事分解成無意義的小單位來研究。肯尼迪認為,這種方法也許特別適用於暢銷文學的創作,但無法帶來對文學和藝術的新認識,因為研究者間主觀層面的微妙差異已經被淹沒在“以計算機為驅動的文學批評”的大數據中了。簡言之,這種數字化的發展於肯尼迪而言是病態的。更可怕的是,在數字化的背景下,這種不安更是難以言表。

肯尼迪不是唯一一位堅決批判數字人文的學者。相較之下,美國文學理論家斯坦利·費什(Stanley Fish)對數字人文的批評更是有過之而無不及。費什曾在《紐約時報》上公開指責:以計算機為導向的人文科學描繪了一種虛假的永恆,這助長了人們對計算機和互聯網的盲從。^②美國詩人、文學評論家亞當·基爾希(Adam Kirsch)也為計算機導致的人文科學的平庸而惋惜。^③丹尼爾·阿靈頓(Daniel Allington)、莎拉·布魯耶特(Sarah Brouillette)和戴維·哥倫比亞(David Golum-

bia)等人^④還擔心,隨着數字和計算機的廣泛運用,人文科學領域也會像硅谷一樣出現對技術的盲信。^⑤邁克爾·哈格納(Michael Hagner)和卡斯帕·赫希(Caspar Hirschi)^⑥等人也批評了數字人文宣稱的對社會知識的誇張性重塑。^⑦2018年,《法蘭克福匯報》更是以《數字人文——繁榮時期的回顧》(“Digitale Geisteswissenschaften - Bilanz einer Gründerzeit”)為題,用了整整兩個版面來介紹人文科學的這一新發展。但文章最終並沒有確定,數字人文究竟是發展的新機遇,還是實證主義的騙局。^⑧

這類針對數字人文的爭論逐年增加。2019年,甚至出現了“數字人文之戰”。學者們在《批評探索》(*Critical Inquiry*)和《高等教育紀事報》(*The Chronicle of Higher Education*)等知名學術刊物上爭論計算機輔助方法對文學專業的價值:計算機輔助方法值得應用嗎?它會不會只能帶來平庸的研究結果?^⑨幾乎不需要多麼敏銳的洞察力,學者們就會發現這場爭論涉及諸多不同的內容:肯尼迪關乎寫作的創意性,費什關乎數字出版路徑,基爾希關乎文學研究理想形態的建構,阿靈頓等人關乎新自由主義,哈格納和赫希關乎對研究中博學美德的稱讚,《法蘭克福匯報》、《批評探索》和《高等教育紀事報》的作者們則致力於捍衛他們所熟悉的人文科學的傳統,以防其可能發生的變化。觀點俗套是這些非議的共同之處:翻來覆去無非是病態、新自由主義或實證主義等說辭,所強調的依然是溫暖、有人情味的文化同冰冷、理性的文明之間的對立。早在1994年,文化學家格奧爾格·波勒貝克(Georg Bollenbeck)就在《教育和文化——德國式闡釋模式的輝煌與不幸》(*Bildung und Kultur. Glanz und Elend eines deutschen Deutungsmusters*)一書中闡明了這一邏輯存在的問題。^⑩

面對數字人文所帶來的不安,人們常常借用文化批評的套話來回避一個關鍵問題:人文科學除了其長期依賴的、熟悉的既定研究形式,能不能或者該不該擁有更多的可能性?人們總能發現各式各樣針對數字人文的批評:現代批評派認為的阻礙作用,文化批評派宣稱的自我毀滅等,這些非議和批評都值得我們認真探討,因為我們能夠從中獲悉人文科學的現狀及其發展的可能性。這場爭論從不同角度涉及了人文科學的自我認知。這篇文章選取了其中的一個方面來探討,即“精確科學”(exakte Wissenschaft)和“不精確科學”(inexakte Wissenschaft)之間的複雜關係。

二、格林兄弟的“準確”與“不準確”

在數字人文遭遇非議的當下,可能並不是仔細權衡各方觀點的好時機。但事實上,關於精確與不精確的人文科學的爭論由來已久。早在1846年的法蘭克福日耳曼學者大會上,雅各布·格林(Jacob Grimm)就已經談到了“不準確科學”(ungenau Wissenschaft)的價值。他指出,“不準確科學”和“準確科學”(genaue Wissenschaft)的概念分別譯自法語“science inexact”和“science exacte”。和我們現在的做法相似,當時的人們也對學科進行了劃分。用雅各布的話來說就是:

準確科學指的是那些所有內容都經過了仔細論證的學科:例如數學、物理和化學。一旦缺乏這種準確性,其所作的全部努力便會統失去效力。不準確科學指的則是那些允許自己在實踐中迷茫的學科,我們所投身的專業就屬於這一範疇。儘管可能長期存在着錯誤和不足,但它們總能從這些錯誤和不足中不斷進步,變得更加純粹:例如歷史研究和語言研究。詩歌本身也是一種不準確的科學。人們幾乎無法要求歷史研究做到完全的準確;陪審團的判決也不是基於算術運算,而是簡單的常識,且常常出錯。^⑪

雅各布補充道:“在戰爭中,炮兵秉持的就是準確、嚴謹的原則,但騎兵在近距離砍殺敵人的時

候,則無需那麼準確。”在他看來,“準確科學”和“不準確科學”之間雖然有所差別,但各有其獨特的價值。例如格林兄弟研究的語言科學,其價值在於:“所有令人歡欣鼓舞的發明創造,……都來自語言描述的創造力。”因此“不準確科學”更接近人的內心。另一方面,“準確科學”也是有價值的,因為它能利用精確性將自然物質解構和重組。雅各布寫道:“所有的杠杆和發明,(以及)那些讓人驚奇或害怕的東西,都是它們的產物。”換句話說:“準確科學”正在改建世界。它們之所以有這種能力,正是因為它們的“仔細論證”。對雅各布來說,“準確”大概指的就是化學的精確實驗,嚴格遵循邏輯的林氏植物分類法,¹²硅藻屬或船蛆屬的新物種,以及數學的嚴謹性。“不準確”描述的則是一組更貼近我們內心的科學。儘管它們可能長期存在着不準確、錯誤乃至缺陷,但顯然,新出現的詞句、歌曲或歷史背景更容易觸動我們。因此,相較於“準確科學”,“不準確科學”為教化國民作出了更大的貢獻。所以直到今天,它們也依然是報刊雜誌上的熱議話題。不難發現,借助“準確”和“不準確”的概念,雅各布·格林描述了兩類科學在研究方法上的不同,同時也涉及雙方在社會影響以及社會地位層面的差異。

從雅各布 1846 年的演講中可以看出,自然科學家的不斷成功着實讓他感到了不安。雅各布認為,自然科學家的能力是“可怕的”:比起非凡的詞句、華麗的辭藻,廣大青年更重視物理學家和化學家的實驗及其應用範圍。他甚至宣稱,大家都在追着物理學跑。當時,雅各布就已經感受到了人們對“準確科學”的傾向性。可見,早在 170 多年前,科學就已經有了向精確化道路發展的趨勢。雅各布也談到了“不準確科學”的“不斷進步”,他說,“儘管可能長期存在着錯誤和不足”,但隨着時間的推移,“不準確科學”也“能從這些錯誤和不足中不斷進步,變得更加純粹”。這裡的“純粹”和“準確”是可以互換的詞彙。在雅各布看來,這種對“準確科學”的傾向性似乎是歷史發展的必然趨勢,他所研究的語言科學,遲早也會被算作是“準確科學”中的一員。

當然,精確的語言科學在格林兄弟時期尚未形成。兄弟二人的詞典編纂項目為語言科學發展成一門精確的科學作出了重大貢獻。在談到兄弟倆的辭典編纂計劃時,威廉·格林說他們想要編纂成一部涵蓋“各個詞彙的自然史”的大詞典。¹³雅各布·格林反對當時普遍強調概念和審美的文學研究,他要求建立以作品的語言構成為基礎的文學史,這導致了 19 世紀及其之後的很長一段時間裡,語言和文學的研究是分開進行的。¹⁴格林兄弟的“德意志研究”為語言科學樹立了嚴謹的工作標準。¹⁵作為“精確性”(Exaktheit)的變體,“嚴謹性”(Streng)能成為他們的指導原則之一絕非偶然。格林兄弟和校勘學家卡爾·拉赫曼(Karl Lachmann)以及語言學家喬治·弗里德里希·貝內克(Georg Friedrich Benecke)一起,將語言科學從以前的信息搜集發展成了嚴謹且專業的獨立學科。

因此,準確性可以說就是格林兄弟的方法論理想。這一方法論將格林兄弟的研究與馮德哈根(Friedrich Heinrich von der Hagen)以及布辛(Johann Gustav Gottlieb Büsching)等日耳曼學者主張的以文化民族或民間教育為取向的語文學研究區別開來。格林兄弟常常尖銳地批評馮德哈根等競爭對手,絲毫不能容忍布辛等人的錯誤和不足。因為格林兄弟認為,語文學的研究也應被嚴格論證。因此,在兄弟二人看來,“準確科學”與“不準確科學”之間的矛盾並非是不可調和的。溫暖的人文科學和冰冷的自然科學的說辭,更多的是人文科學對外的一種策略,其目的是為自己所屬的學科創造合法性。正是因為有了嚴謹的方法論理想,格林兄弟才被算作是語文學的創始人之一。他們把語文學發展成了一門具有共同研究對象的學科,規定了語文學學者應具備的基礎知識,定義了語文學的研究問題,同時也指出了適用的研究方法。除此之外,他們還確立了語文學的學者結構,

據此可以判斷一位學者到底屬不屬於語文學家。按照格林兄弟的標準,馮德哈根和布辛的研究並不屬於語文學的研究範疇。正是得益於格林兄弟的精確傳統,古典語文學家維拉莫維茨(Ulrich von Wilamowitz-Moellendorff)才能在1921年完成了《古典學的歷史》(*Geschichte der Philologie*)一書,這本書的核心便是語文學在語法研究中所遵循的精確性準則。^⑩對格林兄弟和維拉莫維茨來說,嚴格的文本批評是為了重現文本在歷史上的重要意義。格林兄弟把語文學視為一門追求精確的學科,即使是身處浪漫主義文學時期,他們的嚴謹性也絲毫不減。

這就是19世紀中葉發生在格林兄弟身上的“準確科學”與“不準確科學”間的爭論。時至今日,“兩種科學群”^⑪的說法仍然以不同的形式普遍存在,相關的討論也和格林兄弟時期一樣充滿矛盾。兩個陣營的學者對數字人文的評價也往往大相逕庭。例如,肯尼迪就敵視地把計算機輔助方法比作是對藝術品的摧殘。對一些學者而言,“精確人文科學”是不經反思的技術主義;對另一些學者而言,“不精確人文科學”不是科學。從雅各布到肯尼迪,人們討論的主題總能被劃分成這兩類。至於“準確人文科學”最終會決定人文科學的研究範式這一觀點,只是在爭論中被順帶提及。“準確人文科學”到底是一種威脅還是光明的前景,完全取決於學者的立場及其所遵循的論證策略。如果仔細審視雅各布將語文學準確性的方法論合法化的做法,就會發現,其實這並不涉及“準確人文科學”和“不準確人文科學”的對立;肯尼迪和其他批判數字人文的學者也只是在為自己的學科和學術利益辯護,並未因此把自己的學科視為不準確的科學。實際上,沒有學者願意在學術研究中被認為不嚴謹或不準確。

只有把“反對精確的人文科學”這種矛盾性的表述同文化批評的傳統相分離,人們才能更清楚地看到,精確和不精確人文科學的研究對象和研究方法的差異是與學科間不同的作用範圍相關聯的。像格林兄弟那樣研究語言自然史的學者,需要用高度概括的檢驗標準,甚至是語言變遷的法則來進行研究。而那些着眼於藝術品的獨特性或反思文學的主觀藝術性的學者,則強調審美經驗和文學詮釋的個性化。雙方都強調了各自研究對象的複雜性,儘管在具體的研究方法上存在差異,但他們都用自己的方法達成了研究目標。所以,哲學家威廉·狄爾泰(Wilhelm Dilthey)所宣稱的“理解性的人文學科應與嚴格規範的學科相對立”,無論在學科體系上還是在歷史上都是不準確的。更準確的說法是,由於學者對研究對象的理解存在差異,衍生出了對“精確人文科學”的不同理解。例如,對單一詩作的闡釋可以被認為是精確的,因為這種解釋只適用於這一個研究對象,且不能被刻意泛化。反之,對語音變化法則的表述也可以被看作是精確的,因為它適用於語言這一多樣的研究對象,因而可以被概括成普遍的語音規則。

可以看出,關於人文科學準確性的爭論早在計算機和互聯網出現之前就已經存在了。但一直以來,爭論所涉及的觀點大多有失偏頗。只有在精確性成為雙方都熟識的常規概念之後,這一爭論才有意義。格林兄弟便是證明“精確人文科學”並非悖論的範例。對格林兄弟而言,精確性不僅是一種熟識的常規概念,更是一個崇高方法論理想。

三、從聖靈中誕生的“精確方法”

“精確人文學科”的理想並非由格林兄弟首創,它的歷史比格林兄弟所處的時代更久遠,甚至要早於近代自然科學的興起。^⑫正如我們在格林兄弟和當前的數字人文之爭那裡所看到的,人們通常認為精確性不適用於人文科學,它只存在於自然科學、生命科學和技術科學之中,精確人文科學就是個悖論。為了證明這一觀點,人們總是會強調科學史上的關聯性:16至17世紀,人們通過觀

察和測量獲得了有關自然界的客觀知識，近代自然科學隨之興起，直到 19 世紀中葉，近代自然科學才最終形成。^①毫無疑問，望遠鏡、顯微鏡等儀器以及知識的數學化，都在當時促進了新科學理想的實現，即對自然現象進行準確的觀察和規範的描述。從那時起，理論、模型以及系統設計的實驗就成了“精確科學”這一概念和現象的一部分。乍看之下，這些似乎正是人文科學所缺乏的東西，人文科學的研究似乎並不要求達到自然科學所要求的客觀性。

然而，正如悉尼大學科學史教授斯蒂芬·高克羅格(Stephen Gaukroger)所發現的那樣，在科學史上，類型化、刻板化地劃分出精確的自然科學和不精確的人文科學的做法幾乎是站不住腳的。他注意到，強調精確觀察的近代科學與基督教信仰的教派化幾乎發生在同一時代。在那個時代，宗教既扮演着主導性的角色，又發揮着破壞性的作用。^②高克羅格的觀點有啟示意義，因為它可以證明，對那些現在常被列為不準確科學的學科來說，向精確化方向發展的理想同樣適用。例如，如今大多被劃歸到了歷史詮釋學範疇的某些學科，比如文獻學、印章學和古文字學等，它們都在充滿宗教色彩的環境中產生，並且使用了精確的研究方法。

不同的教派和修會都想通過研究證明自己才是對的，強調其他人堅持的都是異端邪說。^③1675 年，耶穌會士帕佩布羅施(Daniel Papebroch)出版了《古文書真偽辨異序》(*Propylaeum antiquarium circa vert et falsi discrimen in vetustis membranis*)一書，第一次詳細闡釋了如何精確區分真假手稿。1681 年，本篤會士讓·馬比榮(Jean Mabillon)發表了開創性的研究成果《古文書學》(*De re diplomatica*)，由此產生了一種對文獻進行精確鑒別的新科學。科學工作應以研究對象的客觀物質性為指導，而非某種先入為主的偏見。只有這樣才能證明，“加爾默羅會^④的宗教傳統來自先知以利亞”或者“羅馬地下墓穴裡的是聖人遺骨”等說法是錯誤的。教士們總在強調自己才具有科學上的準確性，並以此證明其它教派的錯誤。這正好符合了那個時代科學界的共同理想，即對研究對象進行準確的觀察。世界上最早的科學團體之一，1603 年於羅馬成立的山貓學會(*Accademia dei Lincei*)，其標志就是視力極佳的猓狽。^⑤著名的物理學家伽利略以及劇作家波爾塔(Giambattista della Porta)都曾是學會的成員。

有史以來最大的版本項目(Editionsprojekt)——記載聖人言行的《諸聖傳記》^⑥(*Acta Sanctorum*)將“精確人文科學”的精神體現得淋漓盡致。因為一旦涉及神聖，幾乎所有內容都會嚴格遵循精確性原則，以便將其與非神聖相分離。但由於這種區分絕非易事，所以《諸聖傳記》的成書才會跨越 300 餘年。這本書對聖人生平和事跡描述的準確性也達到了史無前例的高度。這一研究所運用的客觀方法，在外交學、印章學或古文字學等領域更是前所未有的。要想分辨出聖人的真偽，就有賴於對手稿的仔細審查，以確保結論的準確和客觀。可見，400 多年來人文科學對研究的精確性一直都有要求。對手稿等文物進行系統且客觀的觀察，這種方法與自然科學所遵循的系統客觀研究法並無二致。

因此，第一部電腦輔助的作品是由神父羅伯托·布薩(Roberto Busa)出版的有關托馬斯·阿奎那(St. Thomas Aquinas, 約 1225~1274)的宗教作品也就不足為奇了。1949 年，布薩神父說服了 IBM 創始人托馬斯·沃森(Thomas J. Watson)幫助創建了包含 1,100 餘萬單詞的《托馬斯著作索引》。布薩的版本繼承了“精確人文科學”所遵循的宗教傳統，同時這也是計算機在文本領域的首次應用。^⑦1965 年，第一篇計算機科學的博士論文由修女瑪麗·肯尼斯·凱勒(Mary Kenneth Keller)^⑧發表，她同時也是“初學者通用符號指令代碼”(BASIC: Beginners' All-purpose Symbolic Instruction Code)的研發者之一。布薩神父和瑪麗修女證明，人文科學也可以是精確的，這一點不言

而喻。

上述歷史告訴我們，“精確人文科學”的歷史比雅各布·格林所設想的還要複雜得多；精確性和客觀性的理想絕不僅限於自然科學。2016年，阿姆斯特丹大學人文學教授任博德(Rens Bod)在《人文科學的新歷史》(*A New History of the Humanities*)一書中表示：對文化規律和文化原則的探索將決定人文科學的歷史；精確的研究方法並非伴隨計算機的產生而出現。從宗教視角出發擺脫宗教偏見，鑒別不同流派的主張，盡可能準確地闡明事實，這些科學理想早在人文科學與自然科學對立之前就已經存在了。因此，自帕佩布羅施和馬比榮時代起，“精確的人文科學”就一直是近現代科學的一部分，而非某種破壞行為。

精確的研究方法對人文科學來說是有意義的，這種方法也總能在人文科學中得到應用。但聖母大學助理教授筭章難(Nan Z. Da)²²和肯尼迪卻否認這一事實。她們認為，這種精確性削弱了人文科學的特殊性和歷史性，因為算法無法理解個體的複雜性。只有整體性的方法才能闡釋文學作品的特殊性。文學是如此的複雜，以至於一切計算都只能片面地把握那些瑣碎的內容。基於這種設想，她們本應該對人文科學複雜的研究對象加以說明，以便支撐他們的觀點。然而，肯尼迪和筭章難都沒有論述，文學的複雜性為何與腦科學或社會學的不同。雖然她們承認，精確的計算機輔助方法可以應用於腦科學和社會學的研究，但她們認為這對文學或史學來說並不適用。同樣具有複雜性，文學和史學的研究對象同生命科學和社會科學的研究對象之間存在哪些區別和聯繫？對此，肯尼迪等人什麼都沒解釋。她們提出的有關文學複雜性的論斷，更多的只是一種沒有依據和理由的個人看法。²³

實際上，在目前針對數字人文的爭論中，肯尼迪和筭章難等人只是延續了前人的舊有觀念——首當其衝的便是威廉·狄爾泰的觀念。狄爾泰曾明確地將人文科學的大部分內容排除在自己的理論之外，例如19世紀被廣泛應用的文本關係學(*Stemmatologie*)，它精確地展現了文本或語言之間的親屬關係。文本關係學的精確性後來影響了20世紀的遺傳學和支序分類學，也就是影響了弗朗西斯·克里克(Francis Crick)、詹姆斯·沃森(James Watson)和羅莎琳德·富蘭克林(Rosalind Franklin)以及他們發現的DNA雙螺旋結構。同樣被狄爾泰排除在外的，還有人們為文學的科學化而作出的巨大努力。例如，恩斯特·埃爾斯特(Ernst Elster)在1897年出版的第一卷《文學研究原理》(*Prinzipien der Literaturwissenschaft*)²⁴中，第一次定義了文學的概念，並為當時的兩門學科(文學和文學史)提供了精確的劃分標準。埃爾斯特將文學研究劃分成了兩個流派：一是他的老師威廉·馮特(Wilhelm Wundt)的實證心理學；²⁵二是新語言學派，以赫爾曼·保羅(Hermann Paul)的《語言史原理》(*Prinzipien der Sprachgeschichte*)和愛德華·西弗斯(Eduard Sievers)的計量語言學²⁶為代表。狄爾泰和他的追隨者們既沒有提到德摩根(August de Morgan)那樣的數學家(德摩根曾嘗試依靠音節和句子的長短來判斷《保羅書信》作者的身份)；也沒有提到門登霍爾(Thomas Corwin Mendenhall)那樣的物理學家(他曾為探究莎士比亞作品的真實作者提出過相關學說)；更沒有提到溫森蒂(Wincenty Lutoslawski)那樣的古典語文學家(他曾致力於用計量分析法確定柏拉圖對話的先後順序)；甚至都沒有注意到俄國數學家馬爾可夫(Andrey Andreyevich Markov)的具有劃時代意義的貢獻(1913年，馬爾可夫借助普希金的詩文小說《葉甫蓋尼·奧涅金》[*Eugen Onegin*]計算出了從給定字母到接下來每一個字母的過渡概率)。眾所周知，“馬爾可夫鏈”²⁷後來在1948年啟發了克勞德·香農(Claude Shannon)，香農發現，通過計算已知詞句在過渡時的平均信息量，人們便可以推定出未知的詞句。²⁸因此，肯尼迪等人所謂的人文科學領域的“獨特複雜性”，

既不成體系，也沒有歷史依據，只是延續了威廉·狄爾泰的觀念。

四、對精確形式的讚揚

對“精確人文科學”的不安在很大程度上與數字人文的形式化特徵有關。這種不安導致了在反對“精確人文科學”的辯論中，特別是日前針對數字人文的爭議中，某些學者對數字人文的誤解。確實，數字人文與詮釋性人文學科的研究方法有所不同。數字人文是對研究問題進行建模，也就是將研究問題形式化，進而利用計算機支持的方法和程序來進行研究。^③

早半個多世紀以前，布薩神父就將他的研究問題，即聖托馬斯的“上帝在這個世界存在”^④指的是什麼，用建立詞表的方式進行了分解和簡化。從研究“上帝存在”到建立“詞語索引”，是一種將研究問題模式化的重要方法，這種方法在詮釋性的人文科學那裡是行不通的。但作為嚴謹的語言學家，布薩神父有充分的理由選用這種方法。他注意到，承載意義的大詞總是緊隨在功能性的小詞之後。比如，“praesentia”在文本中會以詞組“in praesentia”的形式出現，而非單獨的“praesentia”。因此，他在《托馬斯著作索引》中對功能詞給予了和意義承載詞一樣的重視，並嘗試統計出托馬斯著作中的所有詞彙。但由於詞彙量大到遠遠超出了個人所能統計的範圍，所以他開始使用電腦。穿孔卡片幫助他準確地計算出了單詞所在的各個組合。通過這種方式，布薩神父在形式上將研究問題分解成了多個精細的步驟，然後再一步步地將研究結果進行重組，這是典型的精確科學才會運用的方法。因為計數也是需要闡釋的，至少對“精確人文科學”而言，兩者都屬於這個範疇。

另一個例子也可以對形式上的建模和闡釋進行補充，即由意大利藝術評論家、醫生和政治家喬瓦尼·莫雷利(Giovanni Morelli)提出的莫雷利方法(Morellische Methode)。這一方法指的是，不根據模糊的大單位，如提香的構圖或拉斐爾的表現形式，而是根據小的細節來理解意大利的藝術作品。^⑤當被問及“什麼構成了提香或拉斐爾的風格”時，莫雷利說，與當時把畫作為一個整體來鑒賞的傳統方法相比，手或耳的繪畫技巧更能提供可靠的信息。要想研究繪畫風格，就需要將風格這一概念進行精細化的分解，使其形成單獨的、可準確觀察的、進而可測量的量。這種方法的前提是，像提香或拉斐爾等人藝術作品的複雜性，並不是不可研究的。相反，莫雷利假設：複雜性可以通過將細節形式化的辦法來進行建模。這與博蘭主義者協會^⑥在17世紀區分手稿時，或布薩神父將聖托馬斯的著作分解為可數單詞時所採用的方法不謀而合。

在莫雷利之後的一百多年，也就是布薩之後的約半個世紀，英國語言文學家約翰·巴羅斯(John Burrows)在其1987年出版的《計算進入批評》(Computation into Criticism)一書中對簡·奧斯汀的小說進行了開創性的研究。與莫雷利和布薩的方法相似，巴羅斯也對文學風格這一複雜的現象進行了建模，並強調了一個很少受文本研究者關注的形式化特徵——詞頻的分布。^⑦隨後，在對不同年代作者進行的一系列研究中，他發現，一些不顯眼的形式特徵，如一位作者會使用哪些詞彙，這些詞彙的使用頻率及其分布，也可以讓研究者獲得類似於“指紋”的風格，並在此基礎上區分作者。^⑧適合區分作者和文本類型的不是句法順序，也不是那些意義承載詞，而是我們經常使用的冠詞和連詞等功能詞。巴羅斯也是最早將系統樹圖，^⑨即用幹系法將基於距離、層級劃分的數據集進行可視化，再次從生物學帶回到語言學中的學者。此外，他還反駁了“文本作者只是社會建構中的一環”的觀點。這一觀點認為，同時代的作家作品都是對當時社會的折射，因此作品風格相似，很難區分。巴羅斯認為，只要細心觀察，懂得用電腦計算，任何人都能學會如何分辨作者的文體風格特徵。

2014年,追隨着巴羅斯的腳步,我們對不同文學史上的作者歸屬問題進行了研究。通過對作者最常用的詞彙進行Z檢驗,我們發現,這種精確的建模方法不僅可以區分不同歷史時期的寫作,也可以區分19世紀男性和女性作者的不同寫作風格,^①甚至還可以解決其它疑難問題,例如某匿名期刊上的文章可能是由一戰期間作為新聞主編的羅伯特·穆西爾(Robert Musil)撰寫的。^②也就是說,它可以回答一些文學研究的基本問題,例如作者身份的歸屬,男性和女性的寫作差異,時代的建構,以及作者的個體特徵,等等。通過詞頻、詞分布的向量化、感情詞或主題的分布等因素對研究問題進行建模,人們可以更好地從歷史維度或根據作家的個體特徵來把握文學作品的複雜性。

文學不是單一的複雜,而是像自然界和文化中的許多現象那樣,有着多重的複雜性。因此,以肯尼迪為代表的對計算機輔助人文研究方法的質疑確實說明了一定問題,但並不是肯尼迪等人所聲稱的問題。對文學研究中的一些基本問題,計算機能夠做出一些貢獻,但它並不是萬能的。對文學研究問題的建模清楚地表明,在文學研究中,我們應嚴謹地討論,或者說必須嚴謹地討論這種研究方法的可能性和合理性。科學不是結論性的真理,而是在最大程度上就研究問題給出合理的解釋。至少從卡爾·波普爾(Karl Popper)開始,人們就在不斷地強調這一觀點:科學既不意味着確定性,也不能嚴格證明什麼,它只是確定一種可能性,並以此來駁回零假設^③預先建立的聯系。因此,科學並不是要證明所謂的備擇假設,^④而是要發現備擇假設之於零假設的優勢。這種在證偽和證實的方法論層面存在的區別既適用於自然科學,也適用於“精確人文科學”。

可見,數字人文並沒有消解語言學和歷史學,而是用量化的方式對其研究方法進行了批判性的拓展。在一些問題上,數字人文對準確性有着嚴格的要求,例如,在情感詞詞典中,應該用哪種情感理論作為詞條的分類標準。這種對情感問題的建模決定了情感分析適不適合把握文學的情感結構,能不能根據作品中的主導情感來區分流派,以及可不可以從較大的文學作品庫中提取出敘事模式。^⑤因此,肯尼迪是對的,情感分析法確實將小說的複雜結構進行了分解;但她同時也錯了,因為她認為這樣的方法不能說明任何文學問題。

我們看到:對人文科學的研究問題進行建模,是數字人文的基礎。這就要求所有隱含的知識都能得到明確的體現,並精確到可以用計算機輔助方法進行研究的程度。比起我們熟悉的“不精確人文科學”,這種研究方式的步驟更為精細。因此,“精確人文科學”往往比較複雜,成本也更高,其研究方法也無法在短時間內快速取得成果。因為人們既要對文本進行建模,又要對迄今為止的語境變化進行建模。比如,要想研究Wattpad[®]之類的社交平台上的年輕人怎樣撰寫文學作品,以及他們喜歡哪些素材和類型,抑或是他們如何評價文學作品,就必須將這些複雜問題拆分成多個不同的步驟:數據提取,基於機器學習對文學評論進行情感分析,等等。這些步驟絕不僅限於對數學算法的簡單利用,而是更多依賴於相關的專業知識。這裡所指的並不是與計算機和互聯網相關的理論或概念,而是與文學以及讀者有關的知識。所以說,計算機輔助的文學研究並不簡單,它不僅需要詮釋和歷史學的知識,更需要建模和統計學的知識。也正是這種不簡單使得某些學者產生了對數字人文的不安。

五、21世紀的人文科學

在科學史上,方法論層面的差異十分普遍,語文學的研究也不例外。語文學以及相關的人文學科應該接受計算機輔助方法作為其研究方法的延伸。但事實卻並非如此。正如肯尼迪等人對數字人文的批評那樣,不同研究領域的概念和方法論都對“精確人文科學”提出了尖銳地批評,但這些

批評往往缺乏客觀的論證。這些批評的原因是多方面的。因為除了自狄爾泰以來人文科學理想主義的狹隘性,數字人文自身也在發展中帶來了一系列問題:新的基礎知識需要被研究者大規模地掌握;新的研究問題及其合法性也需要得到確認;新的研究方法更需要在人文學科的實踐中得到檢驗;最後,還需要建立不一樣的學術成員結構。這些新問題似乎一下子就出現在了人文科學面前。數字人文也因此成了 21 世紀人文科學的任務和功能之爭的熱議話題。對數字人文的不安是可以理解的,因為專業之所以成為學科,正是因其具有系統上的穩定性:它既不會在所有構成層面上同時受到挑戰,也不會遭遇顛覆性的重構。

對數字人文的尖銳批評還有一個原因,那就是在學科已經較好地實現了它的內外部學術功能的時候,幾乎就沒有理由再進行深刻的變革了。人文學科早已成了“現代社會以科學的形式獲取自身知識的場所”。^①數字人文等精確的人文科學所代表的精細化方法論,因依賴於長期的系統實驗、精確統計以及小組研究,作用範圍是有限的。因此,它們的研究成果更多的是以學術刊物的形式,而非書籍著作的形式出版。此外,數字人文研究大部分需要花費大量的時間,且成本較高,其結論也無法在社會系統中得到直接應用。如貝多芬如何擴大了和聲連接的可能空間?英語動詞形式在近幾個世紀的發展有哪些規律?用統計學方法所得出的此類問題的答案並不是社會所急需的。^②因為針對這類問題,音樂學可以根據現有形式做出更準確的回答,並給出更符合聽眾的音樂評判。語言發展的進化模型也在生物學等領域找到了自己的受眾。進化並不屬於社會自我認知的範疇。這就是為什麼這類研究通常只出現在 *Nature* 或 *PLoS ONE* 等期刊上,而且它們的結論很少會被社會公眾討論。

因此,雖然數字化要求人文科學從根本上改變既有的自我認知,但已經發展完善的人文科學似乎並沒有理由這樣做。與之相對的是文化數據卻呈現出迅速增長的趨勢。圖書電子版本、網絡門戶和語料庫,以及近年來不斷豐富的音樂下載、在線閱讀或訪問博物館的實時數據,它們都在以前所未有的規模和深度被收集並創造着新的文化信息。^③人文科學將這些信息和數據拱手相讓,給了那些互聯網公司,這無異於是停止了以科學的形式從社會獲取知識。數據密集型科學並不是一種宿命,但如果因此而忽視了新數據洪流所帶來的機遇,將會對人文科學的發展帶來不利影響。^④其它一些因素,例如人文科學學生人數的逐漸縮減,也宣告着人文科學的沉寂與消亡,^⑤這就要求我們在另一個歷史性、系統性的層面上探討人文學科的精確性之爭。這篇文章就是要說明為什麼人文科學可以有一個更好的研究基礎和更好的發展前景。“精確人文科學”是有歷史、成體系的科學。“精確人文科學”能否在未來扮演更重要的角色,取決於學者們現在對它的判斷。人文科學沿着精確化的方向發展並不是什麼新鮮事,在人文科學研究中走更精確的新道路也並不是一種破壞性行為。因此,“精確人文科學”對人文科學的進一步發展是有價值的,就算遭受到“拿魚打畫”之類的非議,我們也得去做這件事。

①Richard Lea, “It’s like hitting a painting with a fish: Can computer analysis tell us anything new about literature?”, *The Guardian* (15. September), 2016, URL: <https://www.theguardian.com/books/2016/sep/15/what-is-the-point-of-cultural-analytics-computers-big-data->

literature.

②Stanley Fish, “Digital Humanities and the Transcending of Mortality”, *New York Times* (9. Januar), 2012, URL: <https://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of->

mortality/? rref = collection% 2Ftimestopic% 2FFish% 2C% 20Stanley&action = click&contentCollection = opinion®ion=stream&module = stream_unit&version = latest&contentPlacement=49&pgtype = collection.

③Adam Kirsch, “Technology Is Taking Over English Departments: The False Promise of Digital Humanities”, *The New Republic* (9. Mai), 2012, URL: <https://newrepublic.com/article/117428/limits- digital- humanities- adam- kirsch>.

④丹尼爾·阿靈頓 (Daniel Allington) 是倫敦大學國王學院高級講師, 莎拉·布魯耶特 (Sarah Brouillette) 是卡爾頓大學副教授, 戴維·哥倫比亞 (David Golumbia) 是弗吉尼亞聯邦大學副教授。——譯者注

⑤Daniel Allington, Sarah Brouillette, David Golumbia, “Neoliberal Tools (and Archives): A Political History of Digital Humanities”, *Los Angeles Review of Books* (1. Mai), 2016, URL: <https://lareviewofbooks.org/article/neo-liberal- tools- archives- political- history- digital- humanities/#>.

⑥蘇黎世聯邦理工學院教授邁克爾·哈格納 (Michael Hagner) 和聖加侖大學教授卡斯帕·赫希 (Caspar Hirschi)。——譯者注

⑦Michael Hagner, Caspar Hirschi, “Editorial”, Nach Feierabend, *Zürcher Jahrbuch für Wissenschaftsgeschichte* 9, 2013, S.7- 11.

⑧Lorraine Daston, Peter Galison: *Objectivity*, Boston: MIT press, 2007; “Digitale Geisteswissenschaft — Bilanz einer Gründerzeit”, *Frankfurter Allgemeine Zeitung* 107 (9. Mai 2018), N3f.

⑨Nan Z. Da, “The Computational Case against Computational Literary Studies”, *Critical Inquiry* 45(3), 2019, pp.601- 639, <https://doi.org/10.1086/702594>; “The Digital Humanities Debacle: Computational methods repeatedly come up short”, *The Chronicle of Higher Education* (27. März), 2019, https://www.chronicle.com/article/The- Digital- Humanities- Debacle/245986?cid = wcontentgrid_41_2.

⑩Georg Bollenbeck, *Bildung und Kultur: Glanz und Elend eines deutschen Deutungsmusters*, Frankfurt/ M, 1994.

⑪Jacob Grimm, “Über den Werth der ungenauen Wissenschaften”, *Kleinere Schriften*, hg. v. Karl Müllenhoff u. Eduard Ippel, Bd. 7, Anhang, Berlin 1884 [1846], S. 563- 566.

⑫由瑞典生物學家卡爾·馮·林奈 (Carl von Linné) 提出。林奈採用階梯等級分類法, 將自然界分為“三界”, 即動物界、植物界和礦物界, 界以下依次是綱、目、屬、種, 實現了分類範疇的統一。——譯者注

⑬Wilhelm Grimm, “Bericht über das Deutsche Wörterbuch”, *Kleine Schriften*, hg. v. Gustav Hinrichs, Bd. 1, Berlin, 1881 [1847], S.513.

⑭Rainer Rosenberg, “Zur Bedeutung von Jacob Grimms Konzeption der philologischen Germanistik für die Entwicklung der Literaturwissenschaft”, *Zeitschrift für Germanistik* 7(1), 1986, S.33- 40.

⑮Lothar Bluhm, “Die Brüder Grimm und die Wissenschaft: Anmerkungen aus philologischer und kulturwissenschaftlicher Sicht”, *Neuphilologische Mitteilungen* 106(4), 2005, S.469- 485.

⑯Ulrich Wilamowitz - Moellendorf, “Geschichte der Philologie”, Alfred Gercke u. Eduard Norden (Hg.), *Einleitung in die Altertumswissenschaft*, Band 1, Heft 1, Leipzig, Berlin, 1921, S.1- 80.

⑰Charles P. Snow, *Two Cultures and the Scientific Revolution*, London, 1959.

⑱James Turner, *Philology: The Forgotten Origins of Modern Humanities*, Princeton, 2014.

⑲Lorraine Daston, Peter Galison, *Objectivity*; “Digitale Geisteswissenschaft — Bilanz einer Gründerzeit”, *Frankfurter Allgemeine Zeitung* 107 (9. Mai), 2018, N3f.

⑳Stephen Gaukroger, *The Emergence of a Scientific Culture. Science and the Shaping of Modernity, 1210- 1685*, Oxford, 2006; *The Collapse of Mechanism and the Rise of Sensibility. Science and the Shaping of Modernity, 1680- 1760*, Oxford, 2010; *The Natural and the Human: Science and the Shaping of Modernity, 1739- 1841*, Oxford, 2016.

㉑Jan Marco Sawilla, *Antiquarianismus, Hagiographie und Historie im 17. Jahrhundert: Zum Werk der Bollandisten, Ein wissenschaftshistorischer Versuch*, Tübingen, 2009.

②又稱“聖衣會”、“迦密會”，天主教托鉢修會之一，12世紀中葉創建於巴勒斯坦的加爾默羅山。——譯者注

③David Freedberg, *The Eye of the Linx: Galileo, his friends, and the beginnings of Modern Natural History*, Chicago/London, 2002.

④一部68卷本的百科全書式文獻，按照基督教節日的順序記載了聖徒的生平，最早由波蘭耶穌會士路斯威德(Heribert Rosweyde)構思並撰寫，實則是一部批判性的傳記。——譯者注

⑤Steven E. Jones, *Roberto Busa, S.J., and the Emergence of Humanities Computing: The Priest and the Punched Cards*, New York, 2016.

⑥Mary Kenneth Keller, *Inductive Inference on Computer Generated Patterns*, Madison, Wisconsin, 1965; Denise Gürer, “Pioneering Women in Computer Science”, *ACM SIGCSE Bulletin* 34 (2), 2002, pp. 175 - 180.

⑦Nan Z Da, “The Digital Humanities Debacle. Computational methods repeatedly come up short”, *The Chronicle of Higher Education* (27. März), 2019, https://www.chronicle.com/article/The-Digital-Humanities-Debacle/245986?cid=wcontentgrid_41_2.

⑧Berenike Herrman, Anne - Sophie Bories, Francesca Frontini, Simone Rebora, Jan Rybicki, “Response by the Special interest Group on Digital Literary Stylistics to Nan Z. Da’s Study”, *Journal of Cultural Analytics* (5), 2019, <https://culturalanalytics.org/2019/05/response-by-the-special-interest-group-on-digital-literary-stylistics-to-nan-z-das-study/>; Fotis Jannidis, “On the Perceived Complexity of Literature: A Response to Nan Z. Da”, *Journal of Cultural Analytics*, 2019, <https://culturalanalytics.org/2019/06/on-the-perceived-complexity-of-literature-a-response-to-nan-z-da/>.

⑨Ernst Elster, *Prinzipien der Literaturwissenschaft*, 1897/1911, Halle.

⑩Wilfried Barner, “Literaturwissenschaft”, Robert Harsch- Niemeyer (Hg.), *Beiträge zur Methodengeschichte der neueren Philologien: Zum 125jährigen Bestehen des Max Niemeyer Verlages*, Tübingen, 1995, S.91- 110.

⑪Karl - Heinz Best, “Eduard Sivers (1850 - 1932)”, *Glottometrics* 18, 2009, S.87- 91.

⑫馬爾可夫鏈(Markov Chain, MC)是概率論和數理統計中具有馬爾可夫性質且存在於離散的指數集和狀態空間內的隨機過程。——譯者注

⑬Peter Grzybek, *Contributions to the Science of Text and Language: Word Length Studies and Related Issues*, Dordrecht, 2006.

⑭Julia Flanders, Fotis Jannidis, (Hg.) *The Shape of Data in Digital Humanities: Modeling Texts and Text- Based Resources*, London, 2018.

⑮托馬斯·阿奎那認為天主的存在是不證自明的，但卻也不是無法證明的。在《神學大全》中他提出了證明天主存在的五個證據。——譯者注

⑯Lermolieff, Ivan [i. e. Giovanni Morelli], *Die Werke italienischer Meister in den Galerien von München, Dresden und Berlin*, Leipzig, 1880.

⑰博蘭主義者協會(Die Gesellschaft der Bollandisten)將羅馬天主教會聖徒的生平事跡以手抄本的形式匯編成批判性的版本，在《聖徒行實》(Acta Sanctorum)一書中發表，並附有歷史批判性的評論。這個名字可以追溯到荷蘭神學家 Johannes Bolland (1596~1665)。其初衷是通過對文本和傳統進行批判性的重新評價，保護神學傳統不被新教和當時的理性主義精神迫害。——譯者注

⑱J. F. Burrows, *Computation into Criticism: A Study of Jane Austen’s Novels and an Experiment in Method*, Oxford University Press, 1987.

⑲John Burrows, “Computers and the Idea of Authorship”, Fotis Jannidis, Gerhard Lauer, Matias Martinez, Simone Winko, (Hg.) *Rückkehr des Autors: Zur Erneuerung eines umstrittenen Begriffs*, Tübingen, 1999, S.167- 181.

⑳系統樹圖是生物學上表示親緣關係的樹狀圖解，能很好地描述數據群與數據群之間的關係，是在統計學數據群分析過程中一種非常直觀的表達方式。——譯者注

㉑Fotis Jannidis, Gerhard Lauer, “Burrow’s Delta and Its Use in German Literary History”, Matt Erlin u. Lynn Tatlock (Hg.), *Distant Readings: Topologies of German*

Culture in the Long Nineteenth Century, Rochester, New York, 2014, pp.29- 54.

⑫ Simone Reborá, Berenike Herrmann, Gerhard Lauer, Massimo Salgaro, “Robert Musil, a War Journal, and Stylometry: Tackling the Issue of Short Texts in Authorship Attribution”, *Digital Scholarship in the Humanities*, 2018, <https://doi.org/10.1093/lc/fqy055>.

⑬在統計學中,零假設(虛無假設)是做統計檢驗時的一類假設。零假設的內容一般是希望證明其錯誤的假設。比如說,在相關性檢驗中,一般會取“兩者之間沒有關聯”作為零假設,而在獨立性檢驗中,一般會取“兩者之間有關聯”作為零假設。——譯者注

⑭備擇假設(對立假設),統計學的基本概念之一,假設檢驗中需要證實的有關總體分布的假設,它包含關於總體分布的一切使原假設不成立的命題。——譯者注

⑮ Jian Gao, Matthew Jockers, John Laudun, Timothy Tangherlini, *A Multiscale Theory of the Dynamical Evolution of Sentiment in Novels*, 2016 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC), DOI: 10.1109/BESC.2016.7804470; Jodie Archer, Matthew Jockers, *The Bestseller Code: Anatomy of the Blockbuster Novel*, 2016, New York; Evgeny Kim, Roman Klinger, “A Survey on Sentiment and Emotion Analysis for Computational Literary Studies”, *arXiv:1808.03137 [cs.CL]*, 2018.

⑯ Wattpad 是一個電子閱讀和寫作平台,是社交出版商兼在線免費電子書制作分享社區。作家寫手將自己的小說免費發布在該平台上,在此過程中,作者可以跟讀者互動,甚至通過讀者的反饋來塑造故事。——譯者注

⑰ Wolfgang Frühwald, Hans Robert Jauß, Reinhart Koselleck, Jürgen Mittelstraß, Burghart Steinwachs, *Geisteswissenschaften heute*, Frankfurt/M, 1991.

⑱ Erez Lieberman, Jean-Baptiste Michel, Joe Jackson,

Tina Tang, Martin Nowak: “Quantifying the Evolutionary Dynamics of Language”, *Nature* 449, 2007, pp.713-716; Fabian Moss, Markus Neuwirth, Daniel Harasim, Rohrmeier Martin, “Statistical characteristics of tonal harmony: A corpus study of Beethoven’s string quartets”, *PLoS ONE* 14(6): e0217242, 2019, <https://doi.org/10.1371/journal.pone.0217242>.

⑲ Niels Brügger, “Digital Humanities in the 21st Century. Digital Material as a Driving Force”, *Digital Humanities Quarterly* 10(2), 2016, <http://www.digitalhumanities.org/dhq/vol/10/3/000256/000256.html>.

⑳ Tony Hey, Stewart Tansley, Kristin Tolle, *The Fourth Paradigm: Data- Intensive Scientific Discovery*, Cambridge, 2009, <https://www.microsoft.com/en-us/research/publication/Fourth-paradigm-data-intensive-scientific-discovery/>.

㉑ Heidi Tworek, “The Real Reason the Humanities are ‘in Crisis’”, *The Atlantic* (18. Dezember), 2013, <https://www.theatlantic.com/education/archive/2013/12/the-real-reason-the-humanities-are-in-crisis/282441/>; Kent Cartwright, Sarah Chinn, Tarshia Stanley, *A Changing Major: The Report of the 2016- 17 ADE Ad Hoc Committee on the English Major*, 2018, <https://www.ade.mla.org/Resources/Reports-and-Other-Resources/A-Changing-Major-The-Report-of-the-2016-17-ADE-Ad-Hoc-Committee-on-the-English-Major>.

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 4046

譯者簡介: 龐娜娜, 山東大學外國語學院副研究員, 博士。濟南 250100

[責任編輯 桑海]