

复数的数字人文 ——比较视野下的中国数字人文

陈 静

摘要: 在全球信息时代,以计算机为代表的数字技术、媒介与相应的一整套方法及知识话语已经渗透进以高等教育、学术研究和出版为代表的知识生产系统中,改变着学术研究的对象、相关知识和合法性边缘,也由此产生了应用于人文研究的、新的数据工具、资源以及相应的机构和体制。这是一场变革,以印刷术及书籍为物质基础和认识论前提而建立起的印刷知识生产体系,正在向以数字技术和电子终端为物质基础和认识论前提的数字知识生产体系转变。“数字人文”成为一种数字时代人文学术知识生产的重要内在驱动力。西方的数字人文发展过程较长,充满了争议,而中国数字人文学界如何在后发的时空框架中以西方已有的经验为参照,发现中国语境中数字人文的核心问题及理论体系,建立中国本土化的数字人文理论框架及实践路径,就成为比较视野下中国数字人文发展探索的核心议题。

关键词: 中国数字人文; 比较研究; 网络基础设施; 跨学科

Abstract: In the global information age, especially the computer-based digital technology, the media and the corresponding set of methods and knowledge discourse have penetrated into the knowledge production system represented by higher education, academic research and publishing, which is changing the object of academic research, the relevant knowledge and its edge of legitimacy, and also produced new data tools, resources and the related institutions for humanities research. This is a change that shows the printing knowledge production system based on the materials basis and epistemology of the printing and books is transformed to the digital knowledge production system based on the materials basis and epistemological premise of digital technology and electronic terminals. “Digital humanities” has become an important internal driving force in the production of humanities academic knowledge in the digital age. In the development of digital humanities in the West, it is full of controversy, and how to use the experience of the DH in the West as a reference to find the core issue of digital humanities in China or Chinese studies, and establish the framework and practical path of the localization of digital humanities theory in China has become a fundamental question of China’s digital human development exploration in the comparative perspective.

Key words: China’s digital humanities; comparative studies; cyber-infrastructure; interdisciplinary

作者简介: 陈静,博士,南京大学艺术学院副教授、高研院数字人文创研中心成员,硕士生导师。研究方向:文化和媒介研究、新媒体、数字人文。电子邮箱: cjchen@nju.

edu.cn。此文是国家社会科学基金重大项目“西方美学经典及其在中国传播接受的比较文献学研究”(项目编号: 17ZDA021) 成果之一。

比较何为? 在一个新兴领域尚未得到全面、深入发展的时候,就对之进行一种回顾性的比较,况且还是在一种颇有人为、武断且不科学色彩的“中西”二元对立的框架之下,是否合适? 其合理性何在? 若回溯数字人文发展、繁荣的这短短十多年来(以数字人文正式改名算起),我们会发现有关数字人文的反思性争论持续不断,而且常常以比较的方式出现。比如在有关数字人文名称的讨论中,就有关于“数字人文”与“人文计算”的比较,讨论前者如何对后者取而代之;在对数字人文的价值质疑中,“数字人文”学者反复被质疑的就是如何能突破“传统人文”所能达到的高度,得其不能;在面对有关“数字人文”是一群野心勃勃的掘金者为攫取资源所打造的幻象之质问时,数字人文学者则一再高呼不会技术的守旧派将在技术的洪流中被社会淘汰。诸如此类的争论,持续不断。这些争论中尽管有很多来自数字人文外部的批评,但也有内部社群因为合法性焦虑而引发的建设性反思。但无论是哪种争论,体现的都是数字人文作为一种基于数字知识生产范式在逐渐渗透到各个领域的时候,也在持续地自我发展,逐渐建立起合法性身份、合理性内涵以及反思性自省。然而,如何在全球范围内评价数字人文这种新的知识生产范式的有效性? 其合法性、合理性是否在全球多样性语境中同样成立? 其结果是否能为“数字人文”带来不一样的批判性反思? 这些都成为本文的一个起点,以比较为策略,在中国语境内来对全球知识生产体系中的数字人文进行审视,对数字人文作为一种新的知识生产方式所具有的决定性作用和巨大潜力进行说明。本文特别提出,在大数据、数字技术、人工智能、双一流建设和“新文科”背景下,数字人文将有助于中国人文社会科学实现从印刷知识型向数字知识型转变,并构建与国际人文学术进行对话的新格局。

一、作为数字知识生产转型期的“数字人文”

在全球信息时代,以计算机为代表的数字技术、媒介与相应的一整套方法及知识话语已经渗透进了以高等教育、学术研究和出版为代表的知识生产系统中,改变着学术研究的对象、相关知识和合法性边缘,也由此产生了应用于人文研究的、新的数据工具、资源以及相应的机构和体制。从微观来看,人文学者除了利用计算机及文本处理软件写作、数据库查找资料、社交平台营造社群以及搜索引擎查询信息外,也都开始使用包括计算、统计、GIS、可视化、多媒体等数字技术和方法,对文本、图像、声音、实体对象等等传统人文科学研究对象进行多角度再现、分析与再创造,生产出基于

信息技术的新型学术成果。而从宏观来看,我们可以将这样一场变革描绘为以印刷术及书籍为物质基础和认识论前提而建立起的印刷知识生产体系正在向以数字技术和电子终端为物质基础和认识论前提的数字知识生产体系转变。西方学界从1960年代开始就对这场知识生产转型进行了一系列的讨论。比如从罗兰·巴特、雅克·德里达(Jacques Derrida)、马克·波斯特(Mark Poster)、斯坦利·费什(Stanley Fish)到乔治·兰道(George Landow)所开启的信息批判理论,尤其是后结构主义文本理论与信息技术的融合反思;如弗德里克·基特勒(Frederich Kittler)、沃尔夫冈·恩斯特(Wolfgang Ernst)、艾瑞克·胡塔莫(Erkki Huhtamo)、西格弗里德·齐林斯基(Siegfried Zielinski)等学者从媒介哲学、媒介文化、新媒体研究和媒介考古学等方面开展的媒介本体、媒介和文化、媒介与社会的关系的论述;亦或是由麦克卢汉(Marshall McLuhan)、杰·戴维·博尔特(Jay David Bolter)、理查德·格鲁辛(Richard Grusin)、列维·曼诺维奇(Lev Manovich)、马克·汉森(Mark Hansen)和N·凯瑟琳·海尔斯(N. Katherine Hayles)所致力的新媒体研究、电子文本研究等。这些理论脉络共同绘制了一个信息时代的知识生产转型期的丰富图谱,其中人文学者如何在信息方式变化了的社会语境中重新构建一个知识系统,如何从信息技术、媒介、新媒体和数字媒体的角度来探寻作为主体的人是如何认知、思考和传播知识等问题成为了新一轮知识系统/体制构建场景中的核心议题。与此同时,自1990年代以来,作为知识存储和传播最主要方式之一的图书馆在大力发展基于数字技术的数字图书馆(digital library)的同时,也开始重视图书馆在科研技术中对学术研究的支持实践。比如在美国高校图书馆开始普遍流行的数字学术(digital scholarship)就凸显了图书馆对多学科的研究人员进行快学科合作这一日益增长的需求的积极回应(筱雪、胡琳悦等 13;周力虹、原源、韩滢莹 82)。而作为数字技术与人文研究结合的典型代表,数字人文可以说纵横贯穿了各个人文研究及数字技术的领域,从个人学者到理论建构,再到技术研发和资源建设。尽管直至今日,“数字人文”都并没有形成一个公认的明晰定义,亦或是特定的某些群体性特征,但“数字人文”已经普遍被接受为一种在一定群体内具有共识性的自我身份认知,一种新的科技发展生态环境中的人文学术发展的新路径,更是一种不可避免的数字生产转型期的范式形态。

当然,这里并不是说,数字人文就能取代传统人文研究。事实上,对数字人文的争论在近10年以来一直络绎不绝。数字人文经由早期的文本批评(textual critics),尤其是电子编辑(electronic editing)与“人文计算”(humanities computing)所开辟的两种范式,形成了一个具有丰富、多元且复杂的开放谱系格局,其内涵与范围都处在一个不断地被定义、被争论、被批判从而再被定义的过程中。随着研究边界的日渐扩张,参与者队伍的日益壮

大、教学和科研机构的迅速增多以及出版成果的井喷式出现,都持续为数字人文的发展提供了机遇和条件,但也同时引起了一系列的争议和讨论。这些论争不仅发生在数字人文外部,也发生在数字人文的内部;既发生在推特、微信这样的社交媒体或者像博客、新闻媒体和开放获取的电子书这样的网站平台上,也发生在讲座、会议这样的学术场合,或者报纸、书籍和期刊等传统出版物中,常常可以看到各个学科的学者们、文博档馆员们、计算机工程师、统计学家、设计师们在文字间唇枪舌战、刀光剑影。这些议题既包括量化方法施用于人文研究的正当性及合理性(Nan Z. Da、Ted Underwood、Richard Jean So、Hoyt Long)、数字人文是否是对科学和算法的迷恋(Timothy Brennan、Eric Weiskott、Alan Liu)、数字人文所具有的学术价值(Stanley Fish),还包括对数字人文从业者的身份确认的疑问(Stephen Ramsay)、数字人文与文化及媒介研究的关系(Alan Liu)、数字人文与种族和后殖民主义的关系(Tata McPherson、Roopika Risam)等。在数字人文的社区内部,还有过这样一个项目,试图讨论数字人文学者到底在干什么。^①结果却引发了持续3年的在线讨论,甚至作为历史文献收录到了印刷文本之中。但所有这些,也都从一个方面揭示了数字人文作为广泛获得关注、接受和质疑的新的学术领域,已经在全世界范围内形成了值得重视的强大力量,并且愈来愈深入地进入到学术生态、知识生产的核心之中,成为一种数字时代人文学术知识生产的重要内在驱动力。

二、参照系: 西方数字人文发展

在以英美为代表的西方学术世界中,数字人文所涉及的议题和领域已经非常广泛。苏珊·霍基(Susan Hockey)、威廉姆·马克卡特(William McCarty)和哈罗德·绍特(Harold Short)等人都曾以英文出版物及学术发表为主,从历史和学科结构的角度揭示“数字人文”涉及多个学科,比如语言学与文学、计算机科学、信息科学与图书馆学、哲学、神学与宗教研究、历史学、心理学、工程学、艺术与人文、文化研究等(McCarty & Short 188);研究议题具有交叉性,比如高瑾就指出,目前数字人文研究呈现出“学科技术应用与回顾、数字人文技术探索、文本分析与作者研究、不同语言的计算语

^① 2009年,一个名为“数字人文项目的一天”(the Day of Digital Humanities project)、旨在聚集全世界数字人文者的社区营造项目在线开放,从“计算人文学者到底在干什么?”的问题开始,引发了在线社群中持续3年的对于“数字人文”定义的论争。正如Melissa Terras, Julianne Nyhan, Edward Vanhoutte在《定义数字人文读本》(Defining Digital Humanities: A Reader)中所说的,这场论争其实在最后都并没有形成一个普适性的定义,更多的是体现了数字人文者们有关“数字人文应该是以及可以是什么”问题的思考范围。但就读本中所选择的部分内容来看,很多内容都涉及到了数据、数字技术、新媒体所带来的挑战,以及传统人文研究及人文研究所面临的转型困境。

言研究4个子分区研究”领域(3)。在数字人文项目运作及机构设置方面,国外高校中的数字人文机构多数是以中心的方式存在,有少数是以研究生教育项目或者系出现。这种中心往往会跨越专业和学科的限定,强调跨学科合作及多学科的参与,且往往学校本身会提供一定的技术人员,尤其是图书馆馆员及相关的技术专家参与其中。而从数据资源、工具及技术开发等“硬件”和“系统”资源来看,因为1990年代以来西方对于人文社会科学的“研究基础设施”(research infrastructure)^①的日益重视(其中也包括了上文提到的数字图书馆和数字学术等),新一代研究基础设施——“网络基础设施”(cyberinfrastructure)在过去10年间引发了一场杰弗里·洛克威尔(Geoffrey Rockwell)称之为“基础设施转向”的革命。与之前的人文基础设施相比,数字人文的网络基础设施更强调融合性、合作性以及新兴的研究形式所具有的分散性特质,同时与科学—技术导向或者驱动的基础设计建设有着非常大的差异。这个差异在一些西方的数字人文网络基础设施建设的成果上有显示,比如在1990年代数字人文“新发展”阶段(霍基语)出现的文本编码倡议(TEI, Text Encoding and Interchange),都是开源编码编辑语言协议,通过开放、协作和分享,使得多种类型的文本编辑与标记在更大的学术范围内得以施用,同时,个人学者又可以根据具体项目的特殊性对TEI标准进行修订和改进,促进了语言的丰富性和更大的适用性。类似的项目还有弗吉尼亚大学的杰米·麦甘恩(Jerome McGann)牵头建设的罗塞蒂档案(The Rossetti Archive)^②和19世纪电子学术网络基础设施项目(Networked Infrastructure for Nineteenth-century Electronic Scholarship)^③,都是从数据资源建设开始,以人文学术研究的方法和规范标准对人文文献进行收集、数字化、清理和分享,大大减少了获取数据的成本,也因此促进了以学者为主的数字人文研究的普及和推广(McGann 60)。近几年来西方学界对于网络基础设施也有了进一步的讨论,其中主要着力点就在于基础设施的建设重点问题。乔安娜·朱可就很明确地指出:

对于开展学术活动的新环境的设计工作不能留给技术人员以及图书馆专业人才。图书馆在规划和构想保存、使用甚至创造学术资源的方面是至关重要的伙伴。有同样作用的还有那些技术专业人员。但如果用盖房子来打比方的话,他们就是建筑师和施工队。创造档案库、分析

^① 在美国多称为“网络基础设施”(cyberinfrastructure),在欧洲和澳大利亚称为“e-科学”(e-science),在加拿大则称为“研究基础设施”(research infrastructure)。

^② 罗塞蒂档案 The Rossetti Archive, <http://www.rossettiarchive.org/>。

^③ 19世纪电子学术网络基础设施项目 Networked Infrastructure for Nineteenth-century Electronic Scholarship, <https://nines.org/>。

工具以及对人文和社会科学中的综合数据(aggregate data)进行统计分析的一系列工作只能由一个汇聚了技术、专业和学术的个人专家团队来执行,[……]建立学术模型是一个智识上的挑战,而不是技术的。^①

朱可所谈到的问题其实是一个究竟谁是数字人文学者、谁又来主导数字人文研究的问题。关于前者,西方数字人文学界已经有诸多讨论,包括上文提到的洛克威尔、朱可在内,还有斯蒂夫·拉姆齐(Stephen Ramsay)、杰佛瑞·施纳普(Jeffrey Schnapp)、托德·普莱斯勒(Todd Presner)、彼得·路勒福尔德(Peter Lunenfeld)等人都发表过颇为激进的宣言,要求数字人文学者必须会写代码,必须建造某种东西的标准,引起了极大争议。他们中的部分人在《数字人文宣言 2.0》中更提出了一种普遍性的数字全球主义的观点,倡导用“数字的工具、技术和媒介”来实现学术话语的结构调整,改变已有的知识系统、生产与传播方式:

印刷不再是一种知识生产及(或)传播的唯一或标准媒介;相反地,印刷发现自身被吸收进了新的、多媒体形态之中;而数字的工具、技术、媒介则深刻地改变着艺术、人文和社会科学中的知识生产与传播。数字人文寻求在世界中发挥首要作用。在这个世界中,是大学,而不再是知识和文化的独家生产者、管理者或者传播者,被号召起来去为当下新兴的公共领域中的学术话语打造天然的数字模式,去树立这些领域中的卓越和创新,去实现全球和地方的即时知识生产、交换和传播网络形式。^②

当然这种对数字人文应当由学者来引导、而数字人文学者必须具备一定技术技能的要求,一方面呈现了研究问题及学术导向在数字人文中的重要性,另一方面也在一定程度上体现了数字知识生产对跨学科合作生产的要求。事实上,从这些年来看,跨学科、多学科协作已经成为数字人文一定程度上的“标配”。很多数字人文的中心和团队都是数据分析人员、图书馆员、计算机专家、设计师及人文学者共同参与,这些人往往要求能在项目框架内对人文研究目标、技术实现方式和跨学科工作流程达成共识。比如曾经主编过数字人文里程碑著作《数字人文指南》(*A Companion to Digital*

^① 详情参见 Drucker, Johanna. “Blind Spots: Humanists must plan their digital future”. *Chronicle of Higher Education* 55: 30 (2009). <http://chronicle.com/free/v55/i30/30b00601.htm>. 又,这是作者自译,本文未标注译者处均为作者译。

^② 详情参见“The Digital Humanities Manifesto 2.0”, *Digital Manifesto Archive*, 2009年6月22日(发表时间)。 <https://www.digitalmanifesto.net/manifestos/17/> 最后访问时间:2018年5月27日。

Humanities) 的苏珊·施赖布曼(Susan Schreibman) 就曾强调过数字人文的跨学科性不仅存在于科研项目中,也在课堂教学甚至是学位项目的课程设置中,而且需要参与者都具备对数字技能的理解和使用能力(2016)。而另一位数字人文学者霍伊特·朗(Hoyt Long)也谈到了类似的观点:成为数字人文学者必须要具备一定的技术能力,开展一些跨学科的合作(2017)。同时,数字人文中的“跨学科”指的不仅仅是人文与技术之间需要开展交流,还指在人文学科内部知识生产以及地方与全球知识生产之间要进行跨学科的交流与合作。瑞贝卡·瑞瑟姆(Roopika Risam)针对有关数字人文近年来在理论和实践方面,数字人文与文化研究、性别研究、阶级研究等方面产生的争论就提出,数字人文的跨学科研究必须突破关于理论与实践之间的二元对立,同时必须将数字人文放置在美国1980到1990年之间的学术背景之下,考察其出现的历史语境和意义,从而实现该学科的多样性实践的真正意义。姜文涛也从文学研究学科史的角度对数字人文学术研究的兴起进行了文献学的考察,指出数字人文的出现与发展与西方近代文学以及作为大学学科体制的文学研究有密切关系(56)。这一点对我们理解数字人文的跨学科有着重要意义。我们在数字人文的发展过程中也必须保持清醒的反思意识,对数字人文产生及发展的社会语境要有所认知,才可能真正地发展出适合中国历史、社会及学术语境的数字人文学术范式。与此同时,我们也必须清楚地意识到,西方数字人文发展的过程充满了争议,而中国数字人文学界,如何在后发的时空框架中以西方已有的经验为参照,发现中国语境中数字人文的核心问题及理论体系,建立中国本土化的数字人文理论框架及实践路径,就成为比较视野下中国数字人文发展探索的核心议题。

三、复数的数字人文

数字人文是单数的还是复数的?这个问题由芝加哥州立大学英语系的凯瑟琳·菲茨帕特里克在2011年提出。她在《高等教育纪事报》(2011年5月8日)上发文,以此问题来引出数字人文因其突出的可见性而从者众多,由此引发了对之如何界定的众多分歧和争议。实际上,无论是前文所谈到的布伦南、费什这样的基于印刷的传统学者,还是像苏真、泰德·安德伍德、霍伊特·朗、艾瑞克·威斯科特和斯蒂夫·拉姆齐这样的数字人文学者,或是像阿兰·刘这样的基于数字但非数字人文的学者,尽管他们关于数字人文的观点不同,但都意识到了信息时代中的数字技术革命对学术界及社会所带来的整体性影响,并在不同程度上承认,学术界必须对此影响作出反应。虽然3条批判的路径不同,但依然可以放在知识生产转型这个过程中来看。凯瑟琳·海尔斯将之总结为在知识生产转型期基于

印刷的学者和基于数字的学者之间一种充满张力的状态(Hayles 33-54) 。其矛盾性是因为基于印刷的学者在印刷传统中所形成的研究方法和思维,使之倾向于忽略数字媒介所带来的差异性;而基于数字的学者要么不断地努力向基于印刷的同事解释其数字项目的重要性、意义和价值,要么干脆放弃了与基于印刷的学者的交流,而采取自己的方式来进行更有意义的交流。由此造成了双方之间的对抗和误解。

西方数字人文所面临的知识生产方式转型及其所带来的张力状态在中国语境中同样存在。尽管这一趋势在“数字人文”这个概念被引入到中国之前就已经出现,但中国数字人文发展却整体处在一个起步阶段(陈静 59)。一方面,中国学界与西方学界一样,对全球信息时代的科学技术发展以及相应的大数据浪潮作出了自觉回应,但另一方面,尽管西方数字人文兴起较早,对中国数字人文的发展有着很大的影响,但我们依然无法以一种西方模式来定义或者描述中国数字人文的发展历程和学术形态。这种差异,一方面是因为人文与科技的互动以及跨学科研究自身在中西方语境中存在着历时性的差异,另一方面是因为科学技术在具体人文学科中的施用,需要根据人文学科自身的资源、问题和学科特点进行具体化、定制化和特殊化。因此在面对如何发展中国数字人文发展的浪潮,如何在保持清醒的自觉反思的同时,能有针对性地发展中国模式的数字人文理论与实践路径,就成为了中国数字人文学界在深入开展数字人文同时所必须面对的核心问题。

首先,我们必须意识到,数字人文在中国的发展具有其先在的特殊性,即文本与语言的特殊性。可以说,在中国或者说中文语境中,开展数字人文研究的最大障碍之一是语言障碍。一是因为中文字符及文本有很多特殊性,而目前很多成熟的数据库、文本数字化的基础手段(比如光学字符识别[OCR])、断词方法等都是基于像英文这样的拉丁字母语言的,因此造成了在中文文本,尤其是基于古代汉语的古典文献处理上的很多困难。具体举例来说,中文字符单字多字皆可成词,连断词这个在英文中并不那么困难的问题在中文文本处理中都是首要问题。同时因为目前已有的方法高度依赖于基于大量人工标注的大规模训练语料库或者专业词典,在发现新词和新句的问题上还都是基于监督学习方法,这些都造成了在用数字手段和计算方法处理中文文本上需要进行专门的基础技术研发(Deng 6154)。这些技术难题也成为了中文语境中的数字人文研究深入开展的困难。此外,另一障碍则体现在主要的计算机基础编程语言都是英语,这就促使学习编程语言,尤其是人文学者学习编程语言本身就有一定的门槛,需要花费一定的时间和精力。举例来说,作为计算机存储符号的二进制形式的Unicode用一个字节就能表示一个英文字符,但表示中文字符则需要3-4个字节。虽然常用汉字不过3500个(《现代汉语常用字表》),但事实上汉字多达近10万,3个字节不足以表达全部汉字,因此中文的超大字符

集是用4个字节来表示的。而目前最常被用的关系型数据库MySQL的UTF-8编码实际上只支持3个字节,导致很多特殊中文字符无法被识别,也无法出现在MySQL支持的中文数据库中。而MySQL直到其软件发布的15年后,即在2010年才发布了一个“utf8mb4”的字符集来解决这个问题。如果像后者这样困难,人文学者或许可以通过个人学习克服,但前者的难度则往往超越了人文学者的专业知识和个人能力的边界,而更加要求不同学科的学者参与。这一方面要求中国的数字人文学者在开展相关研究中,必须意识到技术上我们与西方数字人文学界有着巨大差异,像前文提到的TEI编码倡议在中国事实上就很难开展。当然这一方面因为TEI并没有真正进入中文文本处理的实践领域,另一方面也是因为TEI一定程度上并不真正适用于中文文本的处理。此外还因为TEI出现及盛行的1990年代中后期和2000年代早期,中国数字人文尚未得到全面发展。但就目前而言,除了计算语言学、计算机科学、图书馆学等偏向基于技术研发的学科,在处理中文文本时会有专门的一些技术方法和软件外,大部分人文学科的学者还尚未能真正掌握专门的文本处理和计算方法及技术。可喜的是,这两年像R语言和Python语言这样的编程语言的流行和普及,也为人文学者开展自主的文本处理研究提供了机会和可能。比如王涛就用Python语言编写代码,对《德意志人文志》这套工具书进行了结构化处理,并进行了基于统计的人文研究,产生了一定影响力的数字史学研究成果(148)。而台湾大学数位人文研究中心的“DocuSky 数位人文学术研究平台”^①也针对人文学者的特殊性和个人研究需要,整合了多种工具,为人文学者进行个人文本的格式转换、标记与建库、探勘与分析以及可视化、GIS整合等学术研究提供了助力。

其次,我们还必须承认,虽然在过去3年间,“数字人文”作为一个特定领域的发展,在中国也得到了越来越多的重视,各种会议、文章和讨论呈井喷之势,却并没有如在西方语境中那样引发巨大的争议。尽管有一些警惕数字人文的文章出现,但目之所及,很多文章都是在为数字人文叫好。理解此现象,我们并不能就简单认为中国语境对于数字人文更为友好或者更为宽容,事实上这更多的是因为数字人文在中国学界并没有像在英美社会中那样“强大”或者“威胁”传统研究范式,故而并没有能引起更大范围内的讨论和争议。尤其是我们将其放置在中国的科学传统之中,就会发现,对于技术的盲目迷信与无知往往也会导致一种无反思性接受。如许煜在“数码化时代科技和人文的契机”中所论及的,如斯蒂格勒所述,技术与文化的互动在欧洲哲学中是被压抑的,但同时也有着一种连贯性。而在中国,因为科技是由外输入的,所以有关现代科技的哲学诠释是缺席的,从而

^① “DocuSky 数位人文学术研究平台”网址: <http://docusky.org.tw/DocuSky/ds-01.home.html>。

致使“中国可以像美国一样加速发展技术,甚至超越美国,因为在中国,技术发展是几乎没有阻力的;在美国,当一种新的技术出现的时候(如克隆)就会有一个伦理小组来制衡它,但这在中国却是另一个故事”。由此,许煜提出,“因为这个特殊的语境,我们必须谨慎地思考这样一个‘人文危机’,而不只是一窝蜂地去追逐数字人文、网络文化等在西方已发展起来的学科”。在此,我并不想对徐煜对于中国科技哲学的评述做过多评述。但我想说的是,他的论述确实在一定程度上说明了中国人文学界对数字技术及相应的问题普遍缺乏批判与反思能力。由此,我在此处指出的是中国及中文语境对于技术的某种“宽容度”恰恰是因为对科技的“不讨论”(许煜)。这种不讨论并非是不论及科技,而是说,对于科技本身的探究并不深入。仅就数字人文领域而言,尽管这几年来相关研究和论文日益增多,但主要还是集中在图书馆及出版等领域。而文学、历史、哲学及艺术等领域中的学者涉及较少,特别是具体参与一些技术性工作,能懂得软件及代码编写的学者更是少之又少。这一点与西方数字人文群体相比,差异比较明显。当然,人才的培养需要一定的时间,但从深层来看,更重要的是人文学者普遍对于科技的无知与无感。尽管从20世纪90年代以来,关于互联网对于学术研究及写作的影响,已经有颇多讨论,例如网络文学、新媒体、数字文化之类的研究题目俨然蔚为大观,但很多时候研究者对于像数据库、文字处理软件、互联网这样的数字技术及相关软件的认知还存在不足,将之看待为一种“中立”的工具或者单纯的数字产品,而对技术内在的深层运作机制、技术原理、逻辑方法、意识形态和伦理问题还缺少认知和批判性反思。尽管这种技术中立论的观点并非存在于某一特定语境,但就对新媒体、互联网和软件的相关研究而言,中文研究成果较之英语世界尚显得比较单一。

数字人文学者就一定对技术有感了么?未必。事实上很多从事数字技术的人未必会对数字技术本身所具有的偏见、意识形态以及伦理问题有足够的认识。数字人文中具有技术优势的学者往往会强调技术的重要性,以至于对某种算法或者模式推崇备至,似乎在建设一个行业壁垒,并没有考虑到人文研究所具有的特殊性,即更高、更快、更有效并非适用于解决人文问题。这种情况也正是本文开始所谈到的一些有关争论的关节点。在这点上,阿兰·刘对西方数字人文学界此类问题的批评在中国也已经存在:

一方面,数字人文学者担心这个领域太工具[……];另一方面,数字人文学者们又担心数字人文不够“工具化”,无法与有些工程领域相比,在其领域中,工具性彰显了“创新”和“建造”。(2015)

中国数字人文学者在适用于人文学科的工具性建造问题上尚未形成蓬勃之势,但也出现了类似的两难问题。数字人文学者有时也会有意识地

回避“工具化”的问题。比如为了获得更多发表机会、更大范围的认同和降低理解难度,数字人文学者往往强调研究对象、算法模型和结果,而对像数字化、数据库技术有效性、数据格式及清理等前期准备过程中大量看似繁琐实则重要的基础性工作一笔带过。然而,数据结果呈现效果往往取决于这部分的工作。

第三,基于中文的数字人文网络基础设施方面已经做出了很多有益且具有影响力的成果,但目前中国数字人文发展中技术导向模式比较明显,而“学术问题意识”尚有所欠缺。

近10年来,很多团队针对大量有学术及商业价值的数据资源被重复建设,而看似不那么有价值或者开发难度过大的资源却长期被忽略的情况,建立了多个资源平台,开放或半开放给公众使用。在其中比较具有代表性的除包弼德亲自领导的哈佛大学的“中国历代人物传记资料库”(China Biographical Database Project)^①外,德龙(Donald Sturgeon)创办的“中国哲学书电子化计划”(Chinese Text Project)^②、法鼓文理学院经营多年的“中华电子佛典”数据库^③、荷兰莱顿大学魏希德(Hilde De Weerd)教授与何浩洋博士研发的线上古典文献阅读、分析工具“码库思”(MARKUS)^④、台湾中研院的时空地理数位人文研究平台、台湾政治大学古籍数位人文研究平台、上海图书馆“中文古籍联合目录及循证平台”^⑤、中华书局“籍合网”^⑥、搜韵网^⑦等等。这些平台都在对文献进行了一定的前期积累的情况下,开发了适合人文学者使用的工具和工作环境,不仅是像之前的传统数据库那样可以查询,还允许用户与数据库进行互动,部分还提供了应用程序接口,使得其他数据库和系统可以方便地调用其中数据,并灵活地以不同的方式呈现。但基于平台的研究和有影响力的成果尚没有大规模出现,人文文本数据有一定的特殊性,在数据规模、体量上,其实并不如科学数据那样庞大。而与此同时,图书馆、博物馆或者档案馆等机构在建立数据库方面具有一定的先天优势,因此,目前中国数字人文的发展主要还是以信息科学、计算机科学、计算语言学、图书馆学等技术方面比较具有优势的学科主导,主要的论文发表还是集聚在这些领域,而向历史、文学、艺术、哲学等传统人文学科渗透得并不深入。但正如朱可所提到

① 详情见“中国历代人物传记资料库”(China Biographical Database Project), <https://projects.iq.harvard.edu/cbdb/home>。

② 详情见“中国哲学书电子化计划”(Chinese Text Project) <https://ctext.org/>。

③ 详情见“中华电子佛典”数据库(CBETA) <http://cbeta.org/>。

④ 详情见“码库思”(MARKUS) <https://dh.chinese-empires.eu/markus/beta/>。

⑤ 详情见“中文古籍联合目录及循证平台” <http://data.library.sh.cn/index>。

⑥ 详情见“籍合网” <http://ancientbooks.cn/>。

⑦ 详情见“搜韵网” <https://sou-yun.cn/>。

的数字人文想要成为一种普遍的、深入的研究范式,成为传统人文研究的有益补充,产生一系列具有同样有效且影响深远的成果,就必须要以研究问题优先。对此,中国学者也有着比较明确的认识,比如史睿曾撰写过系列文章,对古籍数字资源存在的问题以及文史学者面临的困难进行了说明,提出古籍数字化建设需要人文学术研究者的积极参与(11; 28)。而数字人文也算得上顺势而为,台湾数字人文先驱、资讯工程系特聘教授项洁就曾经描述过从数据库建设向人文研究转向的必然性:

1995年我开始规划并执行台湾大学的台湾史料与藏品的数位典藏工作,这也是在技术上,将新的科技媒体与传统类型史料结合的开始。这项工作进行10年后,我们累积了相当数量的高品质的数位史料,但是我的不安也越来越深。我开始思考,到底如何才能运用资讯科技,在庞大的数位史料基础上从事历史学的学术研究。闭门造车10年多后,才发现在国际已经隐隐约约有一个类似的学问浮现,这就是“数位人文”。近10年来,我找到了越来越多志同道合的朋友,大家均是被数位人文所隐含的可能性深深吸引,也做了不少相关的研究工作。

从此描述也可以看到,数字人文从资源建设转向问题研究是一种必然趋势。要形成中国数字人文的研究方式,就必须从人文学科发展的自身理论资源出发,从一种方法的借鉴和价值的判断回到学理建构。尤其是中国目前正处于数字人文的蓬勃发展期,正在从以数字化方式进行人文资源抢救、整理为主转向以数据驱动的学术研究主导,因此在这个时期必须更好地实现资源整合,并能最大程度地支持人文学术研究,以问题导向来驱动、指导、参与数字资源建设,这样才能对中国数字人文发展的未来大有裨益。

第四,亟需提倡“数字人文”框架下的跨学科合作。跨学科合作之于数字人文研究的核心作用,西方学者已经讨论了很多,虽然在这个问题上,西方数字人文界也存在一定的难度,但就跨学科研究中国发展的历史经验而言,中国数字人文想要突破学科边界,难度更大。一方面是因为目前大学的院系本身是以学科为主,所有的考核和成果都强调学科导向,跨学科成果的认定存在困难;另一方面则是因为资源的分配和各类扶持政策多以专业或者院系为主,而跨学科研究项目较难获得资助。回顾中国数字资源的建设,数据库建设多为国家或者商业经费,参与学者则主要是集中在图书馆、情报及出版专业,人文学者参与不多或者说并不占主导地位,主要原因之一就是数字资源的建设必须要求团队作战,而文史哲学者则更习惯于单打独斗。其次则是因为图情及出版专业有着先天的学科优势,数字图书馆与数字学术在图情及出版专业都有着较长的历史,且成果斐然。再则就是因为长期以来人文学科单个项目的资金有限,而数据资源的建设耗时

久、代价高、劳动密集,单个学者或者小规模团队往往力不能逮。然而,西方数字人文学界所面临的一些问题,却是作为后发的中国数字人文学界或许能避免的,尤其是在目前国家大力提倡“互联网+”“人工智能”等政策的前提下,是否可能借势推动科学技术与人文的结合,从而实现数字人文跨学科快速、顺利地发展,也未必没有可能。在当前建设“新文科”的倡导之下,利用政策红利,来呼应新文科战略的“求变”策略,有望实现科学技术与人文社会科学的深度融合,以科学的求真性、可验性和可重复性来推动人文科学研究的科学化发展,同时以基于数字技术及人工智能的新技术、新方法和新工具来回答新时期人文社会科学问题,并拓展人文社会科学在新时代的研究广度和深度。与此同时,数字人文的跨学科路径突破还有可能在创新人才培养方面实现。尤其是新时代创新型人才培养要求兼顾多学科面向,在具备基础的人文社会学科、自然科学素养、艺术素养的同时,还要兼具数字素养、批判思维和创新素养。数字人文强调理论性与实践性的结合,要求以创新思维突破学科边界来解决与研究社会问题,因此也强调多种理论、能力与技术的融会贯通;以问题为导向来选择工具及方法、以创新为方向来寻求发展路径。在复合型人才的培养上,数字人文强调“数字”所指涉的研究范围,不仅包括使用编程语言进行文本计算、数据库搭建,也包括利用和开发软件开展相关研究,这就要求数字人文实践者不仅需要会使用一定的科技方法来进行人文知识研究、知识传播与管理,还需要介入科学技术的实际研发中去,参与到更具人文素养和人文精神的科学技术发展实践中去。可以说,面对新一轮科技革命与“新文科”发展需求,数字人文无疑指出了一条可能且可行的发展路径,而数字人文相关会议和课程的日益增加也体现了学界的一种自觉性响应,如何进一步推动并将数字人文与“新文科”发展战略实现深度的融合,以更高的战略性角度来思考数字人文在中国的发展模式与路径,建设具有中国特色的数字人文,从而更好地助力于“新文科”战略在理论体系创新和专业改革中的实施力度与深度,成为了数字人文发展的重要命题和时代任务。

引用作品【Works Cited】

Borgman, Christine L. *Scholarship in the Digital Age: Information, Infrastructure and the Internet*. Cambridge: MIT Press, 2007.

陈静 “当下中国‘数字人文’研究状况及意义”,《山东社会科学》7(2018): 59-63。

[Chen, Jing. “The Current Situation of Research on the ‘Digital Humanities’ in Contemporary China”(dang xia zhong guo shu zi ren wen yan jiu zhuang kuang ji yi yi). *Shandong Social Sciences* 7(2018): 59-63.]

陈小荷、冯敏萱、徐润华等《先秦文献信息处理》。北京:世界图书出版社,2013年。

[Chen, Xiaohe, Feng Minxuan and Xu Runhua et al. *Information Process of Pre-Qin Documents* (xian qin wen xian xin xi chu li). Beijing: World Book Publishing House, 2013.]

程佳羽、史睿 “古籍数字资源的知识库建设解析”,《数字图书馆论坛》12(2006): 1-11。

- [Cheng, Jiayu and Shi Rui. "An Analysis about the Construction of Knowledge Based on Digital Resources of Rare-Books" (gu ji shu zi zi yuan de zhi shi ku jian she jie xi). *Digital Library Forum* 12(2006) : 1-11.]
- Deng, Ke. "On the Unsupervised Analysis of Domain-Specific Chinese Texts". *Proceedings of the National Academy of Sciences* 113.22 (2016) : 6154-6159.
- 高瑾 "数字人文学科结构研究的回顾与探索", 《图书馆论坛》1(2017) : 1-9。
- [Gao, Jin. "Review and Study of Digital Humanities Intellectual Structure" (shu zi ren wen xue ke jie gou yan jiu de hui gu yu tan suo). *Library Tribune* 1(2017) : 1-9.]
- Hagles, N. Katherine. *How We Think: Digital Media and Contemporary Technogenesis*. Chicago: University of Chicago Press, 2012.
- Hockey, Susan. "The History of Humanities Computing." *A Companion to Digital Humanities*. Eds. Susan Schreibman, Ray Siemens, John Unsworth. Oxford: Blackwell, 2004. <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-1&toc.depth=1&toc.id=ss1-2-1&brand=default>.
- 姜文涛 "数字人文学科研究与近代文学的兴起——以北美文学研究学科史为中心的文献学考察", 《文化研究》33(2018) : 186-203.
- [Jiang, Wentao. "On the Rise of Digital Humanities in Literary and Cultural Studies — A Philological Exploration Centering on the Disciplinary History of Literary Studies in North America" (shu zi ren wen wen xue yan jiu yu jin dai wen xue de xing qi — yi bei mei wen xue yan jiu xue ke shi wei zhong xin de wen xian xue kao cha). *Cultural Studies* 33(2018) : 186-203.]
- 霍伊特·朗 "一位青年数字人文学者的自白", 《零壹 Lab》(2019-01-17)。 <https://mp.weixin.qq.com/s/F1mewcWNzT0Sw6iWfUPUrg>。
- [Lang, Howett. "A Confess of Young Digital Humanities Scholar" (yi wei qing nian shu zi ren wen xue zhe de zi bai). *01 Lab* (01 - 17 - 2019)。 <https://mp.weixin.qq.com/s/F1mewcWNzT0Sw6iWfUPUrg>.]
- Liu, Alan. *Where is the Cultural Criticism on Digital Humanities?. Debates in the Digital Humanities*. Chicago: University of Minnesota Press, 2015. <https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities>.
- McCarty W, Short H. "Mapping the Field". *The European Association for Digital Humanities* (26-05-2016)。 <http://eadh.org/node/188>.
- McGann, Jerome. "Information Technology and the Troubled Humanities". *Defining Digital Humanities*. Eds. Melissa Julianne, Nyhan Terras and Edward Vanhoutte. Surrey: Ashgate, 2013. 49-65.
- Ramsay, Stephen & Geoffrey Rockwell. "Developing Things: Notes toward an Epistemology of Building in the Digital Humanities". *Debates in the Digital Humanities*. Chicago: University of Minnesota Press, 2015. <https://dhdebates.gc.cuny.edu/projects/debates-in-the-digital-humanities>.
- Risam, Roopika. "Beyond the Margins. Intersectionality and the Digital Humanities". *Digital Humanities Quarterly* 9-2(2015) : <http://digitalhumanities.org:8081/dhq/vol/9/2/000208/000208.html>.
- Rockwell, Geoffrey. "As Transparent as Infrastructure: On the Research of Cyberinfrastructure in the Humanities". *Online Humanities Scholarship: The Shape of Things to Come*. Ed. Jerome McGann. Houston: Rice University Press, 2010. http://rup.rice.edu/cnx_content/shape/m34315.html.
- 苏珊·施赖布曼 "如果你习惯在某个固定领域做研究,那最好别选数字人文", 《零壹 Lab》(2016-12-06)。 <https://mp.weixin.qq.com/s/Y6XcR82MxWmD-qLDI4-0kg>。

- [Schreibman, Susan. "Don't Choose Digital Humanities, If You Used to Stay at the Comfortable Zone for Research" (ru guo ni xi guan dai zai mou ge gu ding ling yu zuo yan jiu , na zui hao bie xuan shu zi ren wen). *01 Lab* (06 - 12 - 2016): <https://mp.weixin.qq.com/s/Y6XeR82MxWmD-qLDI4-0kg>.]
- 史睿 “论中国古籍的数字化与人文学术研究”, 《北京图书馆馆刊》(2) 1999: 28-35。
- [Shi, Rui. "On the Digitization and Humanities Studies on the Chinese Classic Books" (lun zhong guo gu ji de shu zi hua yu ren wen xue shu yan jiu). *Journal of the National Library of China* (2) 1999: 28-35.]
- Svensson, Patrik. "From Optical Fiber to Conceptual Cyberinfrastructure". *Digital Humanities Quarterly* (1) 2011. <http://digitalhumanities.org/dhq/vol/5/1/000090/000090.html>.
- Terras, Melissa. "Disciplined: Using Educational Studies to Analyze 'Humanities Computing'". *Defining Digital Humanities*. Eds. Melissa Terras, Julianne Nyhan and Edward Vanhoutte. Surrey: Ashgate, 2013. 67-96.
- 王涛 “数字人文框架下‘德意志人物志’的群像描绘与类型分析”, 《历史研究》5(2018): 148-166。
- [Wang, Tao. "The Group Portraits and Typological Analysis of German Prosopography (*Allgemeine Deutsche Biographie*) within a Digital Humanities Framework" (shu zi ren wen kuang jia xia de yi zhi ren wu zhi de qun xiang miao hui yu lei xing fen xi). *History Studies* 5 (2018): 148-166.]
- 项洁 “一个台湾数位人文学者的贺词”, 《零壹 Lab》, 2016-10-10。 <https://mp.weixin.qq.com/s/Y6XeR82MxWmD-qLDI4-0kg>。
- [Xiang, Jie. "Congratulations from a Digital Humanities' researcher" (yi ge tai wan shu wei ren wen xue zhe de he ci). *01 Lab* (10 - 10 - 2016): <https://mp.weixin.qq.com/s/Y6XeR82MxWmD-qLDI4-0kg>.]
- 筱雪 胡琳悦 董晓晴等 “基于文献计量分析法的我国数字馆研究进展”, 《天津科技》2(2018): 13-19。
- [Xiao, Xue, Hu linyue, Dong Xiaoping, et al. "Research Progress of Digital Libraries in China Based on Bibliometrics Analyzing Method" (ji yu wen xian ji liang fen xi fa de wo guo shu zi guan yan jiu jin zhan). *Tianjin Science & Technology* 2(2018): 13-19.]
- 徐晨飞 赵文娟 “我国数字人文研究领域作者合著网络分析”, 《图书馆论坛》2019-03-27。 <http://kns.cnki.net/kcms/detail/44.1306.G2.20190619.1319.004.html>。
- [Xu, Chenfei and Zhao Wenjuan. "An Analysis of the Co-authorship Network of Digital Humanities Research in China" (wo guo shu zi ren wen yan jiu ling yu zuo zhe he zhu wang luo fen xi). *Library Tribune* 2019-03-27. <http://kns.cnki.net/kcms/detail/44.1306.G2.20190619.1319.004.html>.]
- 许煜 “数碼化时代科技和人文的契机”, 《文化纵横》2018-11-18。 <http://www.21ber.com/project-post/mulu-5/>。
- [Xu, Yu. "The Meeting Point of Sci-Tech and Humanities in Digital Era" (shu ma hua shi dai ke ji he ren wen de qi ji). *Beijing Cultural Review* 2018-11-18. <http://www.21ber.com/project-post/mulu-5/>.]
- 周力虹 原源 韩莹莹 “中美顶尖高校图书馆数字学术服务对比研究”, 《图书与情报》2(2018): 82-91。
- [Zhou, Lihong, Yuan Yuan and Han Yingying. "A Comparative Analysis of Digital Scholarship Services in China's and the United States' University Libraries" (zhong mei ding jian gao xiao tu shu guan shu zi xue shu fu wu dui bi yan jiu). *Library and Information* 2(2018): 82-91.]