

数字人文与文学研究

庞玉厚

摘要: 随着计算机、互联网和信息技术的发展,知识生产和传播的方式发生了深刻变化。传统的、以纸质媒介为载体的人文学科面临着新的挑战,以计算机、数字媒介为基础的“数字人文”作为新兴的人文学科领域,日益受到瞩目。数字人文为人文学科的发展注入了新的活力,为探索新方法、提出新问题和获得新发现提供了契机。然而,学界对数字人文的学科地位却一直争论不休。同时,数字人文也受到传统人文学者的激烈批评。本文主要考察数字人文的内涵、起源与发展、研究领域和实践活动及其在文学研究中的应用,反思数字人文所面临的机遇和挑战,希望对数字化时代的文学研究和新文科建设有所启示。

关键词: 数字人文; 文学研究; 量化分析; 计算方法; 数字技术

中图分类号: I106 **文献标识码:** A **文章编号:** 1672-0962(2024)02-0104-10

引言

20世纪90年代以来,计算机、互联网、信息技术和数字媒介的飞速发展,极大地改变了知识生产和传播的方式。人们可以对海量数据信息进行存储、检索和分析,这对以纸质媒介为载体、以个体阅读和阐释为主要方法的传统人文学科提出了新的挑战。在此背景下,“数字人文”(Digital Humanities)作为新兴的人文学科领域,受到越来越多的人文学者的青睐。数字人文,简单地说,就是以计算机为基础的计算技术与人文科学相结合而形成的交叉领域。这一术语的正式确立与《数字人文指南》(*A Companion to Digital Humanities*, 2004)一书的出版是分不开的。

基金项目: 中央高校基本科研业务费人文社科项目“数智技术驱动的跨文化外语教学模式研究”(H22XKRCW00010)

作者简介: 庞玉厚,博士,北京交通大学语言与传播学院副教授,研究方向为文体学、叙事学、数字人文。

根据约翰·昂斯沃斯(John Unsworth)的叙述,该名称是2001年底与出版社讨论书名时,由他本人最终提议的(Kirschenbaum, 2012: 5)。在此之前,“人文计算”(Humanities Computing)是这一领域广为熟悉的名字,指对人类的各种表达方式,如文本、音乐、戏剧、绘画等的“自动化分析”,是一种“人文主义”活动(Busa, 2004: xvi)。关于名称的使用,威拉德·麦卡蒂(Willard McCarty, 2013: 118)曾慨叹道:“在2006年,‘人文计算’与‘数字人文’还比肩并存;但从那时起,后者最终胜出了”。显然,数字人文这一称呼更具包容性,更能彰显人文学科的本体地位。

数字人文的起源与发展

在人文学科中运用统计方法进行定量分析的历史十分久远。计量文体学的历史可以追溯到19世纪中叶的语文学传统。1851年,数学家奥古斯都·德·摩根(Augustus de Morgan)在一封书信中提出以定量方法分析词长,并以此为手段考察了保罗书信的作者问题。物理学家托马斯·门登霍尔(Thomas Mendenhall, 1887, 1901)也曾用词频统计方法比较狄更斯和萨克雷小说风格的差异,他还用同样方法分析了莎士比亚、马洛、培根以及其他作家的写作风格,以判断莎士比亚作品的作者身份。英语文学教授卢修斯·舍曼(Lucius Sherman, 1893)也曾使用定量分析方法考察英语散文和诗歌的风格变化。此外,西方计量史学也萌芽于19世纪80年代,20世纪上半叶法国年鉴学派更是为此做出了突出贡献。但是,人们一般将数字人文的历史与计算机的使用联系在一起。1949年,意大利神父罗伯特·布萨(Roberto Busa, 1980)试图借助当时的计算机为托马斯·阿奎那(Thomas Aquinas)以及其他相关作者共约1100万中世纪拉丁语词汇的作品编制词语索引表(index verborum),他得到了IBM公司创始人托马斯·沃森(Thomas Watson)的帮助。鉴于其开创性工作,布萨常被尊为数字人文的先驱。

苏珊·霍基(Susan Hockey, 2004)将数字人文(即人文计算)的历史分为四个发展阶段:起步、联合、新发展和互联网时代。从1949年到70年代初期是起步阶段。许多研究者借助计算机编写索引(concordance)程序,为古代语言、诗歌文本等制作索引表。有些研究机构开始安装计算机设备,以辅助词典编纂、文学资料建档等。有些学者还尝试利用计算机进行词频统计,研究作品风格或鉴定作品归属。这一时期,数据分析的主要对象是文本和数字,靠手工输入到穿孔卡或纸带上,工作量十分繁重。一些研究机构和团队逐渐形成,如剑桥大学于1963年成立了“文学与语言计算中心”(Centre for Literary and Linguistic Computing)。数字人文学者开始组织人文计算方面的会议,并创办专业刊物。

从 20 世纪 70 年代初期至 80 年代中期是第二个阶段——数字人文走向联合。技术人员对索引程序不断进行改进,出现了越来越多的通用软件。基于索引程序的词汇研究为文体分析和语言应用提供了条件。为避免重复劳动,人们在文本数据库的建设和维护等方面开始联合起来。有些学者合作建立专门文本资源库,如希腊语言宝库(TLG)、古英语词典语料库(DOE)、牛津文本档案库(OTA)、法语资源项目(ARTFL)等。数据存储方式也逐渐从磁带转向磁盘,出现了各种数据库处理技术和工具。同时,更多人文计算中心建立起来,大学开始开设关于编程语言、软件操作和应用方面的人文计算课程。数字人文的两个重要国际组织“文学与语言计算协会”(Association for Literary and Linguistic Computing,简称 ALLC)和“计算机与人文科学协会”(Association for Computers and the Humanities,简称 ACH)于 1973 和 1978 年相继成立,定期在英国和北美举办人文计算国际会议。

从 20 世纪 80 年代中期到 90 年代初期,数字人文进入第三个阶段:新发展阶段。个人计算机成为学术研究的重要工具,出现了功能出色的文本分析程序。数字人文的各个领域都得到了进一步发展,出现了更多项目,如帕卡德人文协会(PHI)的古典拉丁语文本数据库和塔夫茨大学的珀尔修斯项目(Perseus Project)。更为重要的是,计算机与人文科学协会、文学与语言计算协会和计算语言学协会(Association for Computational Linguistics)1987 年联合发起了《文本编码倡议 TEI》(*Text Encoding Initiative*),并于 1994 年发布了《电子文本编码和交换指南》(*Guidelines for Electronic Text Encoding and Interchange*)。这套标准被广泛应用于各类在线文本的编码,有力地促进了数字资源的共享。

从 20 世纪 90 年代初开始,随着互联网、特别是万维网的出现,数字人文步入第四个阶段,即互联网时代。研究者可以通过互联网便捷地传输资料,电子学术版本成为讨论的议题。电子资源逐渐多媒体化,包括图像、音频、视频等。许多新项目投入到电子资源数据库的建设中,其中文学、艺术方面的数字档案库有:罗塞蒂档案(Rossetti Archive)、狄金森电子档案(Dickinson Electronic Archives)、奥兰多项目(Orlando Project)、布莱克档案(William Blake Archive)、惠特曼档案(Walt Whitman Archive)、麦尔维尔电子图书馆(Melville Electronic Library)等。越来越多的大学成立了人文计算中心,开设人文计算方面的本科、硕士学位课程。

苏珊·霍基的文章发表于 2004 年,距今已过去将近 20 年,数字人文领域已发生了巨大变化。我们认为,如果以 2004 年《数字人文指南》的出版为分水岭,数字人文已迈入第五个阶段,即全面、深入、飞速发展的数字化时代。2016 年《数字人文新指南》(*A New Companion to Digital Humanities*)的出版也印证了这一点,该书前言说“数字人文应该被视为一个‘独立的学科’还是一套相互关联的方法依然存在争议;但是,时至 2015 年,毋庸置疑,它已是一个充满活力、快速

发展的研究领域”(Schreibman, et al., 2016: xvii)。随着多媒体和数字技术的发展,数据分析的对象从以文字为主向多模态发展,研究者借助新的技术工具开展文本、图像、音频、视频等多种媒介的分析。算法和机器学习在数据挖掘和数据分析中扮演着越来越重要的角色。数字人文的研究领域不断扩展和深化,越来越多的项目获得资助,各种大型数据库、专题语料库和专业网站建立起来。公共图书馆、博物馆、档案馆、出版社以及其他商业机构也积极投入到数字图书馆、数字博物馆和数字档案馆等资源平台的开发以及网络学术出版中。

世界各国的高等院校和科研院所日益重视数字人文的发展,纷纷组建专业团队,成立专门研究和教学机构,提供各种层次的学位课程。欧美著名大学,如牛津大学、伦敦大学学院、斯坦福大学、麻省理工学院、普林斯顿大学、耶鲁大学等,都设立了数字人文中心,伦敦国王学院甚至成立了数字人文系。创始于2007年的centerNet,目前已发展成一个由全球200余家数字人文中心组成的国际化合作和交流网络平台。这些中心在数字人文研究和教学、技术工具研发以及数字资源建设等方面发挥着重要作用。值得一提的是,文学与语言计算协会、计算机与人文科学协会和加拿大数字人文学会(Society for Digital Humanities/Société pour l'étude des médias interactifs)于2006年联合成立了“数字人文组织联盟”(Alliance of Digital Humanities Organizations,简称ADHO)。该联盟拥有自己的开放电子学术期刊《数字人文季刊》(*Digital Humanities Quarterly*),读者可以在期刊网站上免费获取全文,联盟还每年召开一次数字人文年度大会。数字人文正进入一个不断优化组合、全面发展的时期。

数字人文的研究领域、实践与应用

数字人文的疆界一直处于动态变化之中,很难清晰、准确地描述其研究领域和实践活动。2001年4月,时任ALLC主席、计算语言学家安东尼奥·赞波利(Antonio Zampolli)在意大利比萨组织召开了一次由多个学科领域的专家、学者参加的会议,旨在为次年7月在德国图宾根大学召开的ALLC和ACH联合年会制定一个人文计算路线图。在大会发言中,大卫·罗贝(David Robey)基于对各分会组的调查,总结汇报了人文计算的现状及未来的发展方向。威拉德·麦卡蒂和哈罗德·肖特(Willard McCarty & Harold Short, 2002)专门为此绘制了一个粗略的人文计算知识图谱,从中我们可以大略窥见数字人文的主要研究领域和内容。

这个图谱共分为三个区域。首先,居于中心的是共同方法论,对各应用学科所共享的计算方法进行了概述,包括主要的数据类型(如:叙事或连续文本、表格式的字母数字数据或数据块、数字、视觉图像、音乐以及其他声音资料)和数据

处理技术(如:文本分析、数据库设计、数字分析、数字成像、音乐信息检索等技术)。此外,还包括共享的通讯技术、超媒体技术、数字图书馆技术等。其次,位于图谱上方的是与人文计算直接互动的学科群,包括文学与语言学研究、哲学研究、神学与宗教研究、历史研究、法律研究、表演研究、物质文化研究、音乐研究等。这些学科群与数据类型和技术方法处于相互联系、密切互动之中。最后,处在图谱下方的是人文计算涉及的知识领域,包括计算机科学、语言学、社会学、历史学、哲学、文学和艺术等学科的相关知识。其中,计算机科学涉及编程、系统与接口设计、人工智能、超媒体和数字图书馆研究、标记技术等专门知识;语言学包括计算语言学、语料库语言学和语言产业等;在文学和艺术方面,则涉及创造性想象力、修辞与设计以及翻译等。这些知识领域构成共同方法论的理论基础,对理解人文计算的本质至关重要。

数字人文研究在不断拓展和更新,新的议题层出不穷,实践活动丰富多样(Schreibman, et al., 2004, 2016)。在理论、方法和技术方面,不断探索新的方法,开发新的工具。具体涉及计算机编程、资源分类方案、数据库设计与管理、电子文档创建与使用、信息检索、数据建模、文本分析、文本挖掘、文本编码、文本多维标记、音视频分析、地理空间分析、时空分析、统计分析、数据可视化、数字物质化、虚拟世界制作等方面的方法和工具。其中,在文本分析方面,目前内容分析的常用方法有主题模型、情感分析、命名实体识别等;风格分析的方法有聚类分析、距离度量、主成分分析等;结构分析的方法有搭配分析、索引行、词性标注等;而关联分析的技术则有模式识别、序列比对等。

无论是数字化本身,还是数据的收集、整理、存储、检索和分析都离不开技术工具的设计和开发。现在,可供数字人文研究者使用的技术工具和软件资源越来越丰富。加利福尼亚大学圣塔芭芭拉分校英文系的艾伦·刘(Alan Liu)整理了一个数字人文工具箱(DH Toychest),包括音频工具、平台数据管理工具、地图绘制工具、(社会)网络分析工具、编程语言工具和资源、文本分析工具和主题建模工具、视频工具、可视化工具、仿真工具等。DiRT(Digital Research Tools Directory)也提供了数字人文研究的常用工具,涵盖数据收集、数据挖掘、数据可视化、图像编辑、文本分析、统计分析等方面的重要软件。此外,TAPoR(Text Analysis Portal for Research)最初主要提供文本分析和文本检索的工具目录,现在TAPoR3.0还提供非文本数据处理工具以及其他常用工具。

数字人文一直重视基础设施和数据资源的建设、保护和传播。人们关注如何设计和出版具有可持续性的项目,数据收集与各类数据资源(如文档、视觉材料、音像媒体、三维物体和艺术品等)的数字化转换,面向用户的人机互动界面设计,数字图书馆建设,网络出版以及学术文化资源保护等。近年来,还讨论大型图集的

可视化,文物的3D建模,“数字原生”学位论文库,博物馆数字文化遗产的沉浸式虚拟环境制作与体验,文本分析、可视化和设计验证等方面的界面设计,数字人文中的物理计算与桌面制作,定制个性化用户内容访问的技术与方法,物联网的技术和文化,数字学术与数字人文期刊发表中的同行评价,以及数字保存、众包、数字人文软件设计局限性等问题。

在实际研究和应用方面,数字人文探讨的内容涉及不同学科领域,研究课题和任务可谓五花八门,如:文体分析与作者身份研究、语料库标注与统计分析、电子学术编辑、计算机辅助文学文本分析、计算机辅助认知文体学研究、专题学术研究数据库建设、印刷学术与数字资源、数字媒介与电影研究、数字媒介时代的多元叙事、预测计算中的主观阐释与时间建模、机器人诗学(创意写作的自动生成)等。如今,数字人文实践活动的范围更为广泛,包括网络文学、标签行动主义(hashtag activism)、平行实境游戏(alternate reality games)、移动创客空间(mobile makerspaces)、电子游戏对数字人文的启示、复古计算(retrocomputing)、数字人文教学与研究中探索式编程的社会文化和认知价值等。有些研究从学术史角度回顾数字人文的全球史、总体史,或者考古学、艺术史、古典学、历史学、词典编纂、语言学、文学研究、音乐、多媒体、表演艺术、哲学、宗教学等领域的专门史。学者们还热衷讨论数字人文的学科定位和未来行动、数字人文中心的现状与问题、数字人文性别研究(如女性主义数字文学史)以及数字人文全球化等问题。数字人文研究正在不断深入,今天的数字人文学者不是“在用计算机工作”,而是“在以数字方式从事人文工作”(Schreibman, et al., 2016: xvii)。人们不仅对人文信息进行计算建模与分析,还对数字技术及其创新潜力的社会影响开展文化研究。

数字人文是一个多学科构成的研究领域,“数字文学研究”(Digital Literary Studies)是其重要的分支之一(Siemens & Schriebman, 2008)。数字文学研究,又称“文学计算”(Literary Computing)、“计算批评”(Computational Criticism)、“计算文学研究”(Computational Literary Studies)或者“量化文学研究”(Quantitative Literary Studies)等。计量文体学、计算文体学和语料库文体学也可以归入这一范畴。数字文学研究主要考察数字时代的文学研究,即数字化和定量分析给文学研究带来的变化,包括在理论和方法上带来的影响以及在文学研究中的应用。

人工智能、计算语言学和语料库语言学的发展为文学研究提供了强大的技术支持。词长、句长、词频、词汇丰富度、韵律分析、关键词分析等已被证明在手稿研究、作者身份识别、文学模式识别、作品主题、风格、体裁以及文学流派分析等方面是行之有效的。随着机器学习的深入,自然语言处理取得了长足进展,命名实体识别、情感分析、主题模型等文本挖掘技术也被应用于文学研究(秦洪武, 2021: 100)。大型文学语料库和各种专门文学档案库为文学研究、特别是文学史和

具体文学作品的解读以及文学教学提供了良好条件。计算机辅助的文学文本分析在考察长时段文学变化规律,如小说结构与社会空间、情感与阶级的关联、小说时间节奏、文学地理空间分布等方面,具有传统文学研究所不具有的优势(Jockers, 2013; Underwood, 2013, 2018, 2020)。此外,斯坦福大学的弗朗科·莫雷蒂及其文学实验室的同事(Moretti, 2005, 2013; Moretti, et al., 2017)开展的“远读”(distant reading)对文学批评,乃至整个数字人文研究都产生了深远影响。

数字人文面临的机遇与挑战

数字人文的发展方兴未艾,前景广阔。与传统人文相比,数字人文具有数字化、大数据、量化分析、跨学科、团队化、项目化和开放性等显著性特征。首先,数字化是数字人文的基础和前提。数字媒介的兴起改变了传统的印刷文化。借助计算机和数字技术,可以将纸媒的图书、报刊、绘画、照片、音像等资料、甚至有历史文化价值的实物进行数字化,建立各种类型的数据库、数字档案库和数字平台。大数据是数字人文的突出特点。从印刷媒体到数字媒体,各种学术文化资源的载体形式发生了根本变化,极大地改变了知识生产和传播的方式。借助互联网和检索工具,人们可以便捷地获取所需信息。数字人文重视资源的开放性,倡导开放获取和知识共享。

大数据为人文学科研究提供了新的可能性。数字人文是“数据驱动的”(data-driven),它基于各种人文数据库和专题语料库,运用计算方法对数据信息进行定量分析,从而揭示人文现象背后所隐藏的文化规律。计算和量化分析是数字人文的核心特征,计算超越学科边界成为数字人文的方法论焦点。在人文社会科学领域涌现出了不少以“计算”为前缀的分支学科,如计算语言学、计算历史学、计算人类学、计算社会学、计算传播学、计算哲学等,“计算”可谓无处不在。

数字人文还具有跨学科性,而且文理交叉。这就要求研究者不仅需要精通某一人文学科的专业知识,而且需要掌握计算机、多媒体、数据挖掘和数据分析等方面的方法和技术。因此,威拉德·麦卡蒂(Willard McCarty, 1999)声称数字人文不是一个“学科”(discipline),而是一个“交叉学科”(interdiscipline)。此外,数字人文还具有团队化、项目化的特点。学术合作和基础设施建设一直是数字人文的两股形塑力量。数字人文研究需要多个学科领域的专家、学者、技术人员和管理者的共同努力,发挥团队合作精神。此外,大型数据库和专题数据库的建设与维护也需要投入大量人力、物力和财力。

当前数字人文的发展也面临着一些困境和挑战。首先,关于数字人文的学科地位,一直存在争议。有人认为数字人文是一门独立的学科,有人认为数字人文

不过是一套研究方法。威廉·帕纳派克(William Pannapacker, 2011)甚至称其为“大帐篷”(Big Tent)。这一比喻不仅体现了数字人文的开放性、包容性和多样性,而且也暗示了数字人文疆域的流动性、杂乱性和不确定性。因此,如何建设和维护好这一学术共同体是一项长期而艰巨的任务。实际上,数字人文至今未发展成为一门独立的学科,而更像一个由多学科的学者共同实践的人文场域。数字人文学者从各自的学科立场出发,形成不同的研究兴趣和课题。如果说他们有什么共同点,那就是他们享有共同的理念、方法和技术工具。与近年来兴起的“公共人文”(Public Humanities)一样,数字人文更像是一种学术潮流或研究范式。所不同的是,公共人文强调人文研究与公众生活的联系和互动,考虑普通读者的需求,而数字人文的技术门槛超越了一般读者的能力范围,甚至许多人文学者也缺乏技术方面的训练。这虽然可以通过跨学科合作解决,但是具备一定的数据科学知识会有利于更有效地提出研究问题,构建统计模型,并进行意义阐释。如何让数字人文的方法和技术更加普及、优化和人性化,是数字人文所面临的紧迫任务。另外,在机构和教学方面,如何开设新的专业方向、设置新的课程也是亟待解决的问题。数字人文的开放性诉求与数字社会的商业化和垄断主义也存在着矛盾。

数字人文也遭到一些传统人文学者的激烈批评和诟病。著名人文学者、文学理论家斯坦利·费什(Stanley Fish, 2012)对数字人文“革命”提出了质疑,认为其“开放性、互动性、无限性、民主性”的承诺不过是一个神话,数字人文对传统人文学术、出版和教育制度的挑战无益于挽救人文学科。小说家、评论家史蒂芬·马尔什(Stephen Marche, 2012)更加明确地反对数字人文,反对计算方法,认为文学不是数据,文学是数据的对立面,文学的数据化抹杀了差异和品味,抹杀了批评的精致和作品接受史。比较文学与文化研究学者提摩西·布伦南(Timothy Brennan, 2017)将矛头对准数字人文的“技术癖”(technophilia),指出技术手段的运用并未产生更多新的阐释,数字人文将“更多信息”(more information)与“更多知识”(more knowledge)混为一谈。笄章难(Nan Z. Da, 2019)也指出计算文学研究在数据性质、统计方法和工具使用上存在着诸多技术、逻辑和概念上的谬误,声称统计工具无力处理文学的复杂性。

结 语

数字人文的发展,机遇与挑战并存。数字人文是数字化时代知识生产和传播方式转型下的产物,为文学研究和新文科建设提供了契机。计算机和信息技术给人文学者带来了极大便利,使他们从繁重的资料处理中解放出来,大大提高了工作效率。互联网和数字媒介的发展还催生了网络文学以及其他媒介、多媒介的

文学形式,改变了文学生产和传播的方式,对文学研究和文学教学影响深远。另一方面,我们也要认识到数字人文的局限性,如数据模型的主观性和偏见,避免算法依赖,警惕工具理性和唯技术论对人文学科的负面影响(刘永强、吴宏宇,2021:88)。数字人文学者需要积极发扬传统人文注重细读、阐释、批判和伦理反思的精神,更好地推动数字人文的健康发展。

参考文献:

- 刘永强,吴宏宇. 2021. 文化技术理论:后人文主义视角下的身体、物质与符号[J]. 广东外语外贸大学学报(4):76-88+158.
- 秦洪武. 2021. 数字人文中的文学话语研究:理论和方法[J]. 中国外语(3):98-105.
- Brennan T. 2017-10-15. The Digital-Humanities Bust: After a Decade of Investment and Hype, What Has the Field Accomplished? Not Much[J/OL]. <https://www.chronicle.com/article/the-digital-humanities-bust/>.
- Busa R. 1980. The Annals of Humanities Computing: The Index Thomisticus[J]. Computers and the Humanities, 14(2):83-90.
- Busa R. 2004. Foreword: Perspectives on the Digital Humanities[C]//Schreibman S, Siemens R, Unsworth J(eds.). A Companion to Digital Humanities. Oxford: Blackwell: xvi-xxi.
- Da N. 2019. The Computational Case against Computational Literary Studies[J]. Critical Inquiry, 45(3):601-630.
- Fish S. 2012-01-09. The Digital Humanities and the Transcending of Mortality[J/OL]. <http://opinionator.blogs.nytimes.com/2012/01/09/the-digital-humanities-and-the-transcending-of-mortality/>.
- Hockey S. 2004. The History of Humanities Computing[C]//Schreibman S, Siemens R, Unsworth J(eds.). A Companion to Digital Humanities. Oxford: Blackwell: 3-19.
- Jockers M. 2013. Macroanalysis: Digital Methods and Literary History[M]. Urbana-Champaign: University of Illinois Press.
- Kirschenbaum M. 2012. What Is Digital Humanities and What's It Doing in English Departments?[C]//Gold M(ed.). Debates in the Digital Humanities. Minneapolis: University of Minnesota Press: 3-11.
- Marche S. 2012-10-28. Literature Is Not Data: Against Digital Humanities[J/OL]. <https://lareviewofbooks.org/article/literature-is-not-data-against-digital-humanities/>.
- McCarty W. 1999-10-22. Humanities Computing as Interdiscipline[J/OL]. <http://iath.virginia.edu/hcs/mccarty.html>.
- McCarty W. 2013. Tree, Turf, Centre, Archipelago—or Wild Acre?[C]//Terrace M, Nyhan J, Vanhoutte E(eds.). Defining Digital Humanities: A Reader. Farnham: Ashgate: 97-118.
- McCarty W, Short H. 2002-04-26. Mapping the Field[J/OL]. <https://eadh.org/mapping-field>.
- Mendenhall T. 1887. The Characteristic Curve of Composition[J]. Science(9):237-249.

- Mendenhall T. 1901. A Mechanical Solution of a Literary Problem[J]. The Popular Science Monthly, 60(7): 97 – 105.
- Moretti F. 2005. Graphs, Maps, Trees[M]. London: Verso.
- Moretti F. 2013. Distant Reading[M]. London: Verso.
- Moretti F (ed.). 2017. Canon/Archive: Studies in Quantitative Formalism from the Stanford Literary Lab[C]. New York: n + 1 Foundation.
- Pannapacker W. 2011 –07 –31. “Big Tent Digital Humanities”, A View from the Edge, Part 1[J/OL]. <https://www.chronicle.com/article/big-tent-digital-humanities-a-view-from-the-edge-part-1/>.
- Schreibman S, Siemens R, Unsworth J(eds.). 2004. A Companion to Digital Humanities [C]. Oxford: Blackwell.
- Schreibman S, Siemens R, Unsworth J(eds.). 2016. A New Companion to Digital Humanities[C]. Chichester: Wiley-Blackwell.
- Sherman A L. 1893. Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry[M]. Boston: Ginn & Co.
- Siemens R, Schriebman S(eds.). 2008. A Companion to Digital Literary Studies[C]. Oxford: Blackwell.
- Underwood T. 2013. Why Literary Periods Mattered[M]. Stanford: Stanford University Press.
- Underwood T. 2018. Why Literary Time Is Measured in Minutes[J]. ELH, 85(2): 341 – 365.
- Underwood T. 2020. Machine Learning and Human Perspective[J]. PMLA, 135(1): 92 – 109.

Digital Humanities and Literary Studies

PANG Yuhou

Abstract: With the development of computer, Internet and information technologies, the production and dissemination of knowledge have undergone profound changes. Traditional print-based humanities are confronting new challenges whereas computer-assisted digital humanities (DH) is receiving increasingly more attention. DH is expected to reinvigorate humanities with its promises of new methods, new questions and new discoveries. However, it is still controversial whether DH is an independent discipline and it is severely criticized by traditional humanists. This study investigates its definitions, origins in development, research areas and practices as well as its challenges and opportunities with particular reference to digital literary studies, which may shed light on literary studies as well as new humanities in the digital age.

Key words: digital humanities; literary studies; quantitative analysis; computation; digital technologies

责任编辑: 孙毅