

数字人文的技术体系与理论结构探讨

刘 炜 叶 鹰

摘 要 数据驱动的研究范式正在人文学科中越来越多地被采用,其在方法论层面呈现出明显的跨学科性和跨界融合现象,被称为数字人文的方法共同体。该共同体由一整套数字人文的技术体系所支撑,正在影响形成中的数字人文的理论结构。数字人文的技术体系主要包括数字化技术、数据管理技术和数据分析技术、可视化技术、VR/AR 技术、机器学习技术等,理论结构涉及结构化或再造人文、数据化或计算人文、可视化或形象人文、拟实化或增强人文、智慧化或扩展人文,应用覆盖资源富集、知识重构、场景重建、增强艺术等。目前,在宏观层面对数字人文整体方法论和理论结构的研究还不多见,这方面的深入探讨将有助于数字人文的整体性发展,并促进跨领域的研究,从而增进数字时代人文学科的繁荣和发展。图 1。表 2。参考文献 22。

关键词 数字人文 数字化 数据科学 可视化 虚拟现实 增强现实

分类号 G250.7

Exploring Technical System and Theoretical Structure of Digital Humanities

Keven W. LIU & Fred Y.YE

ABSTRACT

The data-driven research paradigm is being applied in the studies of humanities. This is forming the methodological community of digital humanities, with showing characteristic interdisciplinarity and interdisciplinary emergence. The methodological community is supported by the technical system of digital humanities, with affecting the theoretical structure of digital humanities. While the digital humanities develop quickly, the technical system and theoretical structure of digital humanities need to establish. For clarifying the issues, the authors suggested a framework to unify both technical and theoretical core.

By reviewing published articles and present views, the authors concluded total characteristics of digital humanities. It is pointed out that the technical system of digital humanities include technologies of digitalization, data and data analysis, visualization, VR/AR, machine learning, and that the theoretical structure is concluded as re-constructing humanities, digitalizing humanities, imaging humanities, augmenting humanities and smart humanities, with applications covering resource collection, knowledge atomization, environment reconstruction and augmented arts.

It is revealed that digital humanities has become an interdisciplinary field, with merging interdisciplinary knowledge in it. As there are few studies on the overall research for the technical system and theoretical structure of digital humanities, this approach could aid to probe into the integrated and interdisciplinary

通信作者:叶鹰,Email:yye@nju.edu.cn,ORCID:0000-0001-9426-934X(Correspondence should be addressed to Fred Y. Ye,Email:yye@nju.edu.cn,ORCID:0000-0001-9426-934X)

studies of digital humanities, with promoting development and flourish of humanities in digital age. 1 fig. 2 tabs. 22 refs.

KEY WORDS

Digital humanities. Digitalization. Data Science. Visualization. Virtual reality. Augmented reality.

0 引言

数字人文(Digital Humanities, DH),源于人文计算(Humanities Computing),是在计算机技术、网络技术、多媒体技术等新兴技术支撑下开展人文研究而形成的新型跨学科研究领域。20世纪60年代以来,在新兴技术发展的强劲支撑下,人文研究的方法发生了重大变化,伴随人文资料的数字化及其网络分享,人文知识的获取、分析、集成和展示所形成的数字化场景,彻底改变了人文学者进行人文资料组织、标引、检索和利用的习惯,以全新格局接续人文研究,并维持一致性和高效性。海量数据,甚至是数据全集或“完全数据”(指 full data 而不是 sampling data)的采用,给人文研究一个从未有过的广阔视野;各类分析挖掘工具的普遍化和低门槛化,使人文研究从过去以资源占有为决定性因素,转变为如今以提出问题的想象力为驱动的“第四范式”^[1]。数字人文正在给人文学术带来颠覆性变化和影响,当前,数字人文研究人员大多同意 Roberto Busa 的论点^[2],即计算学科的主要作用并不是加速人文学科的进步,而是为人文学科领域研究中长期存在的问题提供新的研究方法,从广度和深度两个维度重构人文研究。

数字人文学科领域的许多学者都致力于将技术融合进学术研究,例如文本分析、地理信息系统(GIS)、互动游戏和多媒体等在哲学、史学、文学、语言学和社会学等学科的应用。作为一个典型的文理交叉领域,数字人文研究人员既有传统人文领域研究者,也有计算机技术和多媒体技术学者,还有从事数据清洗、处理、建模、分析、看护和保存的数据专家。在这三类人员的协作下,诸如数字仓储、内容分析、文本挖掘、

多媒体出版、数字图书馆、信息可视化、虚拟现实、地理信息系统甚至机器学习等多种信息技术已在人文领域得到应用,这也给图书馆服务和图书馆学研究带来了新的“增强”时机。随着原生数字(Born-digital)人文资源的增加,现代人文学科正像自然科学领域业已发生的那样,快速地向可计算化方向发展,变成数据驱动型研究,在此过程中,诸如网络民族志、网络文化、虚拟社会等主题正在成为新的人文研究课题,进而不断地丰富或改写数字人文本身的定义,因此确立数字人文的技术体系和理论结构非常必要和重要。

1 数字人文发展概述

早期的人文计算发端于文学和语言学领域。20世纪90年代初期,人文计算作为一个独立的交叉学科逐渐成形。90年代后期,人文学者开始越来越频繁和深入地使用各种新型的数字技术处理人文资料,并进行人文知识生产,如创作电子文本、扫描古籍图书、使用虚拟技术复原古建筑模型和历史上的都市,开发和建设各种在线的音视频数据库等。人文知识的可计算性快速从文学及语言学领域向历史、音乐、艺术等多个领域扩展。在这一进程中,原生数字资源越来越多,如数字地图、计算机图像、在线网页、虚拟人物等,它们的产生极大地丰富了人文研究的对象,同时也对数字人文的定义及发展产生了深刻影响。1999年,Willard McCarty 讨论了人文计算作为一个独立学科与其他学科的关系^[3],以及如何在制度和学术层面来为人文计算活动提供切实的保障和支撑。从当今专门著作看,数字人文研究的学术共同体已经相当清晰^[4-5]。

国际上,数字人文研究联盟也已出现,欧美发达国家在很多大学内建立了跨学科的数字人文研究中心,各种数字人文主题会议在全球频繁召开^[5]。其中,较为知名的研究中心如美国斯坦福大学的计算机辅助人文研究中心和斯坦福人文实验室、MIT 的 Hyper studio、南加利福尼亚大学的数字人文研究中心、马里兰大学的人文技术研究所、英国伦敦国王学院的人文计算研究中心、日本立命馆大学的京都数字文艺研究中心等;成立了计算与人文联合会(The Association for Computers in the Humanities, ACH)、欧洲数字人文协会(European Association for Digital Humanities, EADH)、加拿大数字人文协会(Canadian Association for Digital Humanities, CADH)、日本数字人文协会(Japanese Association for Digital Humanities, JADH)等,并联合组建了国际数字人文组织联盟(The Alliance of Digital Humanities Organizations, ADHO)。目前,该联盟每年召开一次数字人文大会并编辑出版 *Literary and Linguistic Computing*、*Digital Humanities Quarterly* 等专业期刊。

国际数字人文研究的蓬勃发展也推动了中国数字人文研究的持续进步,继 2011 年国内第一个数字人文研究中心落户武汉大学后,2014 年“数字人文与语义技术”研讨会在上海图书馆召开,2016 年 5 月北京大学主办了首届“数字人文论坛”^[6]。近年,除引介剖析西方及国内的数字人文研究外^[7-10],自主创新研究也正在兴起,诸如古籍文本可视化分析^[11]、中国历史地理数据应用^[12]等,并扩展至整个数字人文研究领域^[13-15]。

2 数字人文的技术体系

由于支持数字人文的技术一直处于不断变化完善过程中且新技术不断涌现,数字人文的方法论基础尚未稳定,典型特征还不明显,应用领域也未形成相对成熟的范式,学术共同体内也缺乏共识。

整体而言,数字人文的方法学基础主要是映射和模拟。映射是指将某一学科的研究对象和历史材料经过数字化之后在虚拟世界里建立起一套相对应的系统;模拟是指新系统是对学科领域真实场景的仿真。仅仅将文献数字化是不够的,还必须建立起与真实世界相一致的概念抽象和逻辑框架,当然仿真的程度可以因问题和目标的不同而不同。这其实就是“基于数据的研究”。

美国著名数字人文专家约翰·安斯沃斯(John Unsworth)提出数字人文活动具有的特性可以用七个“学术原语(Scholarly primitives)”来概括^[16],分别是“发现(Discovering)、注释(Annotating)、比较(Comparing)、参考(Referring)、抽样(Sampling)、说明(Illustrating)和表示(Representing)”,可以理解成数字人文常用的七种基本方法。该理论后来被欧洲学者 Tobias Blanke 和 Sheila Anderson 等规范为五个基本原语^[17-18]:发现(Discovering)、收集(Collecting)、比较(Comparing)、发布(Delivering)和协作(Collaborating),以及多个“二级原语”。这种归纳其实是 Willard McCarty 和 Harold Short 于 2002 年总结的数字人文研究“方法共同体(Methodological Commons)”的抽象化和理论化^[19]。

将这些抽象活动与数字人文研究的生命周期过程结合起来,可以清晰地了解在数字人文研究的不同阶段,信息技术在各类活动中是如何起作用的(参见图 1 和表 1)。由此可以进一步明确,信息技术平台作为数字人文研究的基础设施,一方面提供所有数字人文研究所需的素材、工具和交流平台,另一方面也必然带来数字人文在宏观上的整合和一体化现象,而在内部形成问题跨界、学科融合和工具共通的领域特征。

2.1 数字化技术:构造数字人文的资源基础

越来越多的图书馆、博物馆、文化机构以及大学科研机构开始进行人文资料的数字化和网络化,建立了大量的人文主题网站、专题数据

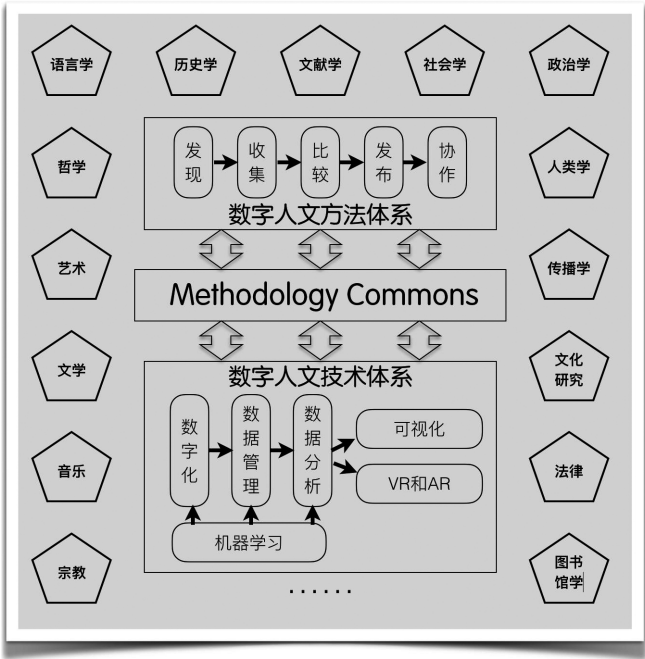


图 1 数字人文技术方法体系构成的方法共同体

表 1 数字人文技术体系的组成

技术类型	具体技术组成(不限于)	实现功能(学术原语)
数字化技术	扫描,拍摄,采样,捕捉,图形设计,3D 建模等技术	收集 Collecting 发现 Discovering
数据管理技术	文本编码,语义描述,本体建模,数据库设计,多媒体搜索,语义搜索,数据看护,名称实体提取(NER),API 数据服务等技术	收集 Collecting 发现 Discovering 协作 Collaborating
数据分析技术	文本分析(词频、共现、关联、向量、概率),聚类分类,主题分析,内容挖掘,时序分析,地理空间分析,社会关系分析等技术	比较 Comparing
可视化技术	信息美学,知识地图,主题图,关联呈现,场景模拟,历史仿真等	比较 Comparing 发布 Delivering
VR/AR 技术	人机交互技术,脑机界面,认知技术,互动测量,游戏化学习,计算机竞技等	发布 Delivering 协作 Collaborating
机器学习技术	自动分类,图像视频音频识别和分析,个性化服务,精准推送,深度学习,超级计算,机器绘画,机器作诗等	发现 Discovering 比较 Comparing 协作 Collaborating

库,并向大众开放。同时,大量的原生数字资源也被纳入这些人类记忆机构的馆藏,作为应对数字化时代挑战的基本策略。例如,许多国家

图书馆都有规模不等的万维网存档项目(Web Archiving),推特公司每年都把它完整记录交由美国国会图书馆保存,还有一些非营利性公

司(如 Internet Archive)有计划地从事互联网存档。这些工作不仅丰富了数字图书馆的内容,也加快了人文学者间的知识共享以及人文知识的大众化普及。支撑数字人文研究是数字图书馆建设重要的目标之一,这也构成了图书馆学关注数字人文的重要原由。在应用驱动下,人文知识的网络数据库正在发展。

2.2 数据管理技术:构造数字人文服务系统

数字化只是实现传统数字人文素材向数字世界的映射,能够被计算机所存储、处理和展示,仅仅转换了一种存在方式而已。要充分发挥数字媒体的作用,还必须进行“数据化”,即按照知识单元的方式来组织领域知识,从而能够构造一个模拟领域应用的知识环境。要做到这一步,通常的“文本化”也是不够的,虽然文本化能够支持基于字符匹配的全文检索,但字符串还不是严格意义上的数据,必须把子串所表达的人名、事物名、时间、地点等析取出来,赋予独立的标识符(正像古希腊哲学家巴门尼德所认为的那样,没有名称的事物是无法独立存在的)。近年来数据科学发展迅猛,已形成一整套数据管理方法论体系,其中本体技术、语义万维网相关技术(如关联数据)等是专门针对知识单元(语义单位)提供解决方案的技术,在数字人文领域已经有了非常成功的应用。

2.3 数据分析技术:构造数字人文应用平台

科研工作除了必须有原始素材之外,还必须有一系列科学方法。基于数据的研究之所以成为一种新型的科研“范式”,主要原因在于其已形成了特色鲜明、行之有效的基于数据的研究方法体系。通常数字人文的研究方法有文本分析(包括词频分析、共现分析、关联关系分析等)、内容(主题)分析、时序分析、空间分析、社会关系分析等,例如数字文学研究就通常要用到文本分析、内容分析,社会关系分析及可视化等,而数字史学在文本分析、内容分析之外,还经常用到时空分析、实体关系分析以及有关历

史人物的社会关系分析等。一个好的数字人文应用平台除了提供海量和规范的数字化信息和数据之外,还应该提供大量的软件应用和统计分析工具等。

2.4 可视化技术:构造数字人文的直观形象

可视化(Visualization)手段也是数字人文平台应该提供的重要服务内容。可视化是利用计算机图形学和图像处理技术,将数据转换成图形或图像在屏幕上显示出来,并进行交互处理的理论、方法和技术。可视化技术的直观操控和实时交互为数据表示、数据处理、决策分析等提供了强有力的支持技术。对于数字人文研究而言,可视化技术能够让复杂庞大的数据形象地展示出来,直观明显地呈现结果。此外,正飞速发展的 VR/AR 技术也需要图形图像可视化技术支撑。

2.5 VR/AR 技术:构造数字人文的交互环境

虚拟现实(Virtual Reality, VR)和增强现实(Augmented Reality, AR)技术,也称灵境技术或人工环境,是利用电脑模拟产生三维空间的虚拟世界,通过用户视觉、听觉、触觉等感官实时、直接观察或操控虚拟空间中的事物,能让用户具有身临其境的感觉。VR/AR 技术集成了计算机图形、计算机仿真、人工智能、传感、显示及网络并行处理等技术,是数字人文技术中的高新技术。VR 和 AR 的结合被称为 MR(Mixed Reality),其最大的可能应用是给数字人文以超越时空的“场景再现”能力,例如在人工智能等技术的辅助下,历史学家可以借助穿戴设备,“穿越到”一定的历史场景中,与“古人”进行随心所欲的对话和交流。

2.6 机器学习技术:构造智能系统实现智慧服务

人工智能的目标是让机器学会做人的工作,而机器学习作为人工智能的核心,目的是让机器学会编程后自动找到解决问题的办法,完

成复杂的任务。机器学习意味着计算机掌握的知识将会远多于人。前不久,以阿尔法狗为代表的深度神经网络算法在下围棋方面取得突破性进展,预示着计算机的认知对人类而言是一个黑箱,这将成为常态,而我们人类只需通过输入输出来控制其算法的改进,使其作为一个“超级大脑”,更好地继续为人类服务。目前机器学习在诸多领域执行专门任务已超越了人类的能力,有望大规模取代人的工作,例如金融分析师、法律文员、汽车驾驶员等。对于数字人文来说,机器学习能够大规模代替人工,进行资料的分类、组织、图像语音识别、名称识别、模式识别、关系发现、跨媒体检索、知识搜索等等,过去我们梦寐以求的自动标引已经基本不成问题。机器学习还能在智能化服务方面发挥独特优势,使数字图书馆的服务更加人性化和个性化,更加精准。人工智能正在改变人类社会的几乎所有方面,也将影响到数字人文的研究。

在上述技术支持下建立的数字人文平台能够对同样的素材从不同角度、不同领域进行研究,各类研究过程和成果又能反过来丰富数字人文平台的数据资料和服务性能,使数字人文系统自身也处在一个不断递归和发展完善的过程中。这正是数字人文技术体系展现出的力量。

3 数字人文的理论结构

数字人文自诞生以来已产生了巨大影响,但也争论不断,由于其从方法论角度颠覆了许多学科固有的基本问题和领域边界,且涉及几乎所有学科,现有的科学学、科学哲学或社会科学哲学等相关理论很难给予恰当的解释,亟需进行理论创新。欧美学者对此虽然已经进行了大量研究,也有数本专著发表,如《数字人文指南》^[5]、《数字人文新指南》^[20]、《数字人文论辩》^[21]等,但该领域还远未定型^[22]。伴随技术体系发展进步,数字人文理论必将在发展中逐步确立和完善。以上述数字人文技术体系为依

托,数字人文呈现出数据技术支持下的结构化、可视技术支持下的可视化、VR/AR 技术支持下的拟实化、机器学习技术等支持下的智慧化趋向,预计会逐步形成其理论结构。

3.1 结构化或再造人文

语言文字是人文表达的基础,数字人文对人文的作用首先体现在增进数字(电子)文本的规范化和标准化程度上。为此,计算与人文联合会(Association for Computers and the Humanities)等已于1987年成立文本编码组(Text Encoding Initiative, TEI),并发布了多个文本编码标准。这些标准主要面向人文社会科学研究,其目的是便于机器编码和识别,提高数字文本的传播、分析和教学效率。目前,TEI 主要由英国牛津大学、美国布朗大学、美国维吉尼亚大学等机构资助,其标准得到数字人文学者的广泛支持,已成为数字人文领域最基本的文本编码标准。这些标准的实施为人文社科领域的文本挖掘提供了极大的便利。

一旦语言文字得到数字文本和结构化数据标准的支持,其“可计算性”得到释放,或将在很大程度上使“再造人文”成为现实。

3.2 可视化或形象人文

可视化技术促成了直观人文研究成果,使人文研究可获得直观形象支持,从而实现对人文学术的视觉直感,由此产生的形象人文体现并表达了人文学术的进步。

这一成就首先在博物学、历史学等领域获得应用,越来越多的学者开始借助地理信息系统技术对历史文物和历史事件进行静态和动态的可视化展示研究,如中国开发的“虚拟故宫”和日本开发的“虚拟京都”皆已上线。美国加州大学圣巴巴拉分校的空间综合社会科学研究中心(CSISS)和英国伦敦大学学院(UCL)的高级空间分析中心(CASA)集聚了地理信息系统、地理学、经济学、物理学、计算机科学等多学科的专家,集中研究社会经济系统在时空演变中的

客观规律以及相应的对策方案,可望系统地推进形象人文。

3.3 拟实化或增强人文

以虚拟现实(VR)、增强现实(AR)等人工智能技术作为方法对人文进行拟实处理,使之具有超文本性和交互性,形成增强人文,其中尤以增强艺术最具魅力。

利用计算机视频捕捉并以运动分析技术进行人体运动轨迹的建模本身即是技术前沿。该技术可以构建人体运动模型,进而为制作虚拟人物提供数据基础,因而已被广泛应用于电影制作以及虚拟场景和虚拟人物制作。用此技术构建的虚拟人物可以达到与真实人物几乎完全一样的运动姿态,这为传统舞蹈、戏剧、电影等在虚拟现实场景中的再现提供了技术支持。

3.4 智慧化或拓展人文

智慧化通常有两方面的含义:内容的智慧化和服务的智慧化。内容的智慧化有赖于对数据进行结构化编码,提供知识的基本结构单元,通过知识本体的构建和关联关系的计算而模拟人脑的认知过程进行前端“机器学习”,而在后端达成一定的“智慧”;服务的智慧化通常是指对固有研究过程和分析模式的模拟,当找到背后的规律之后,能够使得多学科的同类问题迎刃而解(将来提出适当的问题可能比解决问题更为困难)。目前大量的数字人文研究最困难的部分基本都是充当数据苦力,辛苦取得一些可视化结果就能够让人“哇~!”,但未来三五年之后再来考察,目前的大多数成果基本上都是“小儿科”。可以想见,当未来的人文学者在从事研究的时候,时刻都有一个掌握了所有数据(或某一领域的完整数据),并了解所有研究方法的机器人助手服务在侧,取得让人眼睛一亮的成果的唯一障碍只能是想象力了;能够提出多大意义的问题,就能够取得多大意义的成果。这将在很大程度上拓展人文研究的领地!我们目前最有意义的工作,就是构建这样的基础

设施。

智慧应该使人文更强大而不是相反。人文学科在本质上有一种“来自于人性,服务于人性”的特殊性,不像纯粹的自然科学大多是客观规律总结,人文学科的宗旨是服务于人类的福祉,因此它特别强调人与人之间的协作、社区功能的营建以及真善美的达成。在机器横行的将来,数字人文将是我们得以立足的根,培育我们在可能与机器产生冲突时具有的终极智慧。

4 应用展现

在上述技术及理论框架下,数字人文已呈现出多姿多彩的风格和特点,表2通过不同领域的有趣案例,展示了数字人文研究在数据采用和方法技术方面的做法,以及它们如何为研究目的服务。这些案例可能还比较肤浅和初步,但我们从中可以看出数字人文作为一个学科领域,信息技术应用的许多相通的地方,可为当前和将来构建统一的平台和基础设施提供参考。

特别地,数字化和数据化导致的结构化与可视化、拟实化结合将共同驱动数字人文在以下应用上充分地展现。

4.1 资源富集

数字图书馆本身就是一类富集数字资源的实体,可为数字人文提供资源基础。正因为如此,所以数字人文的兴起受到图书馆学界的广泛关注和积极策应。然而,数字人文对数字资源的要求还会向深度延展,比如深入到语言文字内容。因此,资源富集在数字人文中将是重要的基础应用。例如,各类语料库是支撑计算语言学和语料库语言学研究的重要资源,目前发展极为迅速,很多研究机构业已建立面向各个领域的大型语言语料库,这正从根本上改变着传统语言学研究的范式。

4.2 知识重构

数千年文明积累已为人类带来丰富多彩的

表 2 数字人文研究常用方法技术案例

领域	案例举例	目的	数据	方法/技术
语言/文学研究	唐诗研究 ^①	用词习惯、人生轨迹、排行榜	唐诗、作者(生卒、籍贯、关系)、论文、编年信息	词频统计、地理关联、主题分析
社会学/关系研究	唐朝诗人关系研究 ^②	发现唐朝诗人之间的关系	人物库(别名)、诗歌文本	引用关系分析、名称实体识别
历史研究	苏东坡的“流浪地图” ^③	绘制苏东坡一生创作活动时空图	生平、地名、作品年代、数字地图、数据转换	时空可视化、地理信息平台
语言/文学研究	当我们读唐诗时我们在读什么	文学作品(唐诗)分析	唐诗文本	语言分析、内容分析、关联分析
文学/艺术研究	1分钟内复习天龙八部	文学作品(小说)分析	天龙八部整部小说	时空分析、内容分析、地图服务、社会关系/活动分析

知识,对知识的保存、组织、检索、利用和重复利用已形成专门化和专业化领域。而今数字人文的蓬勃兴起和广泛延展,将提供重构知识的技能和机遇,进一步激活并再生人类知识,犹如飞行时代的再次地理大发现:在人类不会飞行的时代,以哥伦布为代表的航海家们曾凭借航海技术环球航行发现新知。而今人类不仅能借助飞机翱翔蓝天,而且能在数字人文技术支持下回顾知识,展翅蓝天能让人类获得与从地平面和海平面看世界不同的视野,数字人文技术再让这种视野如虎添翼,知识重构必将在其中发生,而重构的知识将会启迪文明进步。

4.3 场景重建

利用虚拟技术进行场景重建也是数字人文的重要应用,例如,虚拟的遗址场景可以给观众带来身临其境的感觉,进而提升考古发现的吸引力。因此,数字人文正在深刻改变着考古学

的面貌。又如在甲骨文研究方面,利用甲骨文信息管理系统,研究者将甲骨文词汇聚类,可以根据断代、祭祀、军事、战争、职官、尤其是地名,调阅有关资料,这给甲骨文学术研究带来很大的便利,把学者从繁琐的资料查阅工作中解放出来,极大地提高了研究效率。

此外,除了利用数码扫描仪对古籍和文物进行二维或三维扫描外,考古学者还可利用数字人文技术进行三维虚拟遗址绘图、文物虚拟复原、色彩还原等工作;艺术学者则可利用类似技术进行艺术设计乃至星空艺术想象。

4.4 增强艺术

通过应用 VR/AR 技术,艺术家们在人机交互环境中进行构思和创作,犹如身在自由的灵幻时空中或在梦境中创造,这赋予艺术过程以新的含义,此即增强艺术。

增强艺术吸引艺术家的特别之处在于,艺

① 参见:http://m.thepaper.cn/newsDetail_forward_1675675?from=groupmessage.

② 参见:http://mp.weixin.qq.com/s?__biz=MzI0NTUxMjgyOA==&mid=2247483750&idx=1&sn=dd883b547a3fc4343a3dcce1abea3719&chksm=e94c2e75de3ba7631ffd7abff8a89ea56fda63b2f3d3bb81fd845ef5fd3e9207b41230900288&scene=21#wechat_redirect.

③ 参见:http://mp.weixin.qq.com/s?__biz=MzI0NTUxMjgyOA==&mid=2247483793&idx=1&sn=63316aa8d9b01c93510512403df1bd6e&chksm=e94c2e82de3ba79450fd3444dae11b053c9a873f48c02f1e0ddd42d5f2c045c8f610a12edc0f&scene=21#wechat_redirect.

术思维与科技工具的密切交融和深层渗透,产生了全新的认知体验,进而可能创作出史无前例的艺术作品。

与传统平面新媒体或多媒体艺术相比,增强艺术在交互性和扩展人机对话方面优势明显而独特。正是通过以新型人机对话为基础的交互性艺术创作过程,艺术家们可以借助软硬件的流畅组合来促进作者与作品之间的互动与交流,将数字世界和真实世界结合在一起,实现参与和操控的有机结合,甚至模拟进入作品内部去观察和操纵作品,因此赋予增强艺术强大的生命力。

而这些应用只是数字人文已实现的部分,尚未实现的理想如大海般深广,有待来人继续去发现。

5 小结与建议

自20世纪90年代初数字人文研究兴起以来,人文学者利用数字技术对人文资料进行处理和存储的例子已屡见不鲜。史学、文学、语言学、艺术学、考古学、民族学等领域都已相继展开各类数字人文研究项目,人文学者与技术学者的跨界合作研究已经成为一种流行趋势。从欧美发达国家的数字人文发展历史来看,数据化、公众科学和智慧化是三个明显的趋势,因而数字人文的建设应该关注三个方面:①建设和开发一批人文数据资源库和基础性应用软件,从数字化到数据化,构建数字人文研究的基础

设施;②发展通用研究方法来理解和分析新的数据集,使研究过程和方法平台化,支持公众科学和开放存取的新研究模式;③大力开展新型的基于数据的研究,以机器学习带动智慧服务,创新人文学术的研究方法和研究范式。

随着我国科学技术研究水平的提升,各类保障平台逐步完善,数字人文学科的发展条件已经具备。为加快我国人文学术创新,建议国家重视数字人文基础设施建设,从以下三方面推动数字人文研究:

(1)设立一批数字人文跨学科研究中心。建议鼓励各高校跨院系成立以传统人文学科为核心的数字人文研究中心,实现文理学科之间的优势互补和资源共享,同时创新管理机制,以利于开展跨学科的数字人文研究。

(2)鼓励我国学者积极参与国际数字人文项目、联盟和会议。建议与国外学者共建数字人文项目,在技术层面通过国际化促进我国数字人文研究的快速发展和持续进步。

(3)加强人才培养和深度科研。建议有基础的一流大学结合教学与科研创建人才培养项目和深度科研项目并融通,必要时可设立数字人文硕士点或博士点。

综上所述,数字人文研究正在兴起,在技术方法的强劲支持和推动下,数字化、数据化、可视化、结构化、智慧化交融,正在生发日新月异的变化,值得及时关注和及早布局,以参与和促进数字人文的创新。

参考文献

- [1] Hey T, Tansley S, Tolle K. 科学研究的第四范式:数据密集型科学发现[M]. 潘教峰, 张晓林, 译. 北京: 科学出版社, 2012. (Hey T, Tansley S, Tolle K. The fourth paradigm: data-intensive scientific discovery[M]. Pan Jiaofeng, Zhang Xiaolin, trans. Beijing: Science Press, 2012.)
- [2] Busa R. The annals of humanities computing; the index thomisticus[J]. Computers and the Humanities, 1980, 14(2): 83-90.
- [3] McCarty W. Humanities computing as interdiscipline: is humanities computing an academic discipline?[C]. Paper delivered at IATH, University of Virginia. 5 November 1999.
- [4] Burdick A, Drucker J, Lunenfeld P, et al. Digital_humanities (open access eBook)[M]. Cambridge: MIT Press, 2012.
- [5] Schreibman S, Siemens R, Unsworth J. A companion to digital humanities[M]. Oxford: Blackwell, 2004.

- [6] 朱本军, 聂华. 跨界与融合: 全球视野下的数字人文——首届北京大学“数字人文论坛”会议综述[J]. 大学图书馆学报, 2016, 34(5): 16-21. (Zhu Benjun, Nie Hua. Crossing boundaries and engaging communities: digital humanities in a global perspective—conference review of the 1st Peking University Digital Humanities Forum [J]. Journal of Academic Libraries, 2016, 34(5): 16-21.)
- [7] 王晓光. “数字人文”的产生、发展与前沿[G]//全国高校社会科学科研管理研究会. 方法创新与哲学社会科学. 武汉: 武汉大学出版社, 2010. (Wang Xiaoguang. The generation, development and frontiers of “Digital humanities” [G]//National Academy Social Sciences Research Management. The methodological innovation and development of Philosophy and Social Sciences. Wuhan: Wuhan University Press, 2010.)
- [8] 戴安德, 姜文涛, 赵薇. 数字人文作为一种方法: 西方研究现状及展望[J]. 山东社会科学, 2016(11): 26-33. (Dai Ande, Jiang Wentao, Zhao Wei. Digital humanities as a method [J]. Shandong Social Sciences, 2016(11): 26-33.)
- [9] 周晨. 国际数字人文研究特征与知识结构分析[J]. 图书馆论坛, 2017(4): 1-8. (Zhou Chen. Analysis of features and knowledge structure of digital humanities research abroad [J]. Library Forum, 2017(4): 1-8.)
- [10] 高胜寒, 赵宇翔, 朱庆华. 国内外数字人文领域研究进展分析[J]. 图书馆杂志, 2016(10): 9-18. (Gao Shenghan, Zhao Yuxiang, Zhu Qinghua. A review on digital humanities at home and abroad [J]. Library Journal, 2016(10): 9-18.)
- [11] 欧阳剑. 面向数字人文研究的大规模古籍文本可视化分析与挖掘[J]. 中国图书馆学报, 2016, 42(2): 66-80. (Ouyang Jian. Visual analysis and exploration of ancient texts for digital humanities research [J]. Journal of Library Science in China, 2016, 42(2): 66-80.)
- [12] 夏翠娟. 中国历史地理数据在图书馆数字人文项目中的开放应用研究[J]. 中国图书馆学报, 2017, 43(2): 40-53. (Xia Cuijuan. The opening and application of Chinese historical geography data in digital humanities projects of libraries [J]. Journal of Library Science in China, 2017, 43(2): 40-53.)
- [13] 刘伟, 谢蓉, 张磊, 等. 面向人文研究的国家数据基础设施建设[J]. 中国图书馆学报, 2016, 42(5): 29-39. (Liu Wei, Xie Rong, Zhang Lei, et al. Towards a national data infrastructure for digital humanities [J]. Journal of Library Science in China, 2016, 42(5): 29-39.)
- [14] 柯平, 宫平. 数字人文研究演化路径与热点领域分析[J]. 中国图书馆学报, 2016, 42(6): 13-30. (Ke Ping, Gong Ping. The evolution path and hot topics of digital humanities research [J]. Journal of Library Science in China, 2016, 42(6): 13-30.)
- [15] 郭金龙, 许鑫. 数字人文中的文本挖掘研究[J]. 大学图书馆学报, 2012, 30(3): 11-18. (Guo Jinlong, Xu Xin. Text mining research in digital humanities [J]. Journal of Academic Libraries, 2012, 30(3): 11-18.)
- [16] Unsworth J. Scholarly Primitives; what methods do humanities researchers have in common, and how might our tools reflect this? [EB/OL]. [2017-03-20]. <http://people.virginia.edu/~jmu2m/Kings.5-00/primitives.html>.
- [17] Blanke T, Hedges M. Scholarly primitives: building institutional infrastructure for humanities e-Science [J]. Future Generation Computer Systems, 2013, 29(2): 654-661.
- [18] Anderson S, Blanke T, Dunn S. Methodological commons: arts and humanities e-Science fundamentals [J]. Philosophical Transactions, 2010, 368(1925): 3779.
- [19] McCarty W, Short H. Mapping the field [C/OL]. [2017-03-20]. Paper given at an ALLC meeting in Pisa, 2002. Available at: <http://www.allc.org/node/188>.
- [20] Schreibman S, Siemens R, Unsworth J. A new companion to digital humanities [M]. Hoboken: Wiley - Blackwell, 2016.
- [21] Gold M K, Klein L F. Debates in the Digital Humanities [M/OL]. Minnesota: University of Minnesota Press, 2016 [2017-04-10]. <http://dhdebates.gc.cuny.edu/>.
- [22] Zeng M L. Smart data for digital humanities [J]. Journal of Data and Information Science, 2017, 2(1): 1-12.

刘 炜 上海图书馆研究员, 副馆长。上海 200031。

叶 鹰 南京大学信息管理学院教授, 博士生导师。江苏 南京 210023。

(收稿日期: 2017-05-15)