

中国数字人文年会

数实共生:预见数字人文未来图景

——2023 年中国数字人文年会综述

何悄吟 王晓光

摘 要 由武汉大学信息管理学院与中国索引学会主办,武汉大学信息管理学院、武汉大学文化遗产智能计算实验室、武汉大学大数据研究院及武汉大学信息资源研究中心承办的第五届中国数字人文年会(CDH2023)于2023年12月9—10日在武汉光谷召开。年会以“数实共生:预见数字人文未来图景”为主题,下设十个专题分论坛,受到了国际国内高等院校和社会业界的广泛关注。文章对年会投稿论文、项目内容做了分类评述,对相关技术路径选择、呈现形式等情况进行了分析与总结,以期达到窥探现阶段中国数字人文研究焦点与进展,探讨中国数字人文研究面临挑战和应对方式,展望未来研究趋势的目的。

关键词 数字人文; 数实共生; 会议综述

分类号 G12; G322

作者简介 何悄吟,武汉大学信息管理学院博士研究生,Email:heqiaoyin126@126.com; 王晓光,武汉大学信息管理学院副院长、教授,武汉大学文化遗产智能计算实验室主任。

0 引言

大数据与人工智能技术的崛起,极大地推进了数字文明的蓬勃发展,给传统文化和文化遗产领域的承继、挖掘与焕新提供了无限广阔的可能性空间。以 ChatGPT 为代表的大语言模型和生成式 AI 技术预示着新一轮内容生产革命的开始。信息边界逐渐消融,不同模态的信息经由深度学习越来越趋向统一表示。在信息操纵能力日益增强的背景下,数字人文研究如何深化而促进科技与人文融合共生、人文与科技相互赋能获得新生,成为传统文化、文化遗产等相关产学研领域面临的现实课题。这既是本届年会主题的由来,也是中国数字人文既往年会议主题的拓展。

主题为“文化遗产数字化与数字人文研究”的第一届文化遗产数字化国际研讨会(2019,敦煌)吹响了我国数字人文研究的集结号(因此也被视为首届中国数字人文年会)。此后,无论是第二届中国数字人文年会的“积淀与超越:数字人文与中华文化”(2020,上海),第三届中国数字人文年会的“时代经纬:迈向新文科的数

字人文”(2021, 南京), 还是第四届中国数字人文年会的“全景与远景: 数字人文与国家文化数字战略”(2022, 北京), 都是论秩有序、学理清明的划时代的宏大主题。武汉大学幸执本届年会, 承续前行主义之议, 赓以问题研究为务, 聚焦“数实共生: 预见数字人文未来图景”, 于2023年12月9—10日在武汉光谷兴会海内外同仁。来自北京、上海、武汉、广州、南京、西安等城市的高等院校、科研院所、产业界及美国、英国高等院校的近400名专家和学者, 贡献了6场主旨报告、10个主题分论坛、27场主题报告和1场数字人文教育圆桌对话, 全景式地展示了现阶段中国数字人文研究的新进展、新突破、新成绩, 探讨了数字人文的新主题、新方法和新趋势。其况之盛, 其果之硕, 承前启后。

本文拟对年会的主旨演讲、录用论文、优秀项目等情况进行综述, 回顾会议概况, 陈罗中国数字人文研究的现状及焦点。

1 主旨演讲: 当下挑战与应对策略

分别来自国内五所高校和美国肯特州立大学的主讲嘉宾就教育、学术研究、科技创新、产业经济等主题做了六场精彩各异、启人之思的主旨演讲。

1.1 数字人文教育与数字人文研究

数字人文是从20世纪末发展起来的利用数字技术辅助社会科学及人文学科问题研究的一种研究范式和研究思维。2023年4月4日教育部公布的2022年度普通高等学校本科专业备案和审批结果表明, 数字人文已正式进入了最新版《普通高等学校本科专业目录》^[1], 成为中国官方正式认定的一个新兴学科, 但其学科体系、学术体系、话语体系的建立都还在起步摸索阶段。中国人民大学国家一级教授、数字人文研究院院长冯惠玲在首场主旨报告《数字人文教育的理想与羁绊》中阐述了国际数字人文教育的发展态势, 并结合中国国情和教育现状, 介绍了中国数字人文教育的学科建制和教学特色, 指出该学科复杂的跨学科特性和由此为高等教育组织机构和学科格局带来的更大潜力和更广泛可能。报告在展望数字人文学科发展的同时, 尤其分析了中国数字人文教育的现实情况和数字人文领域发展可能带来的学科格局变化对培养目标、课程结构、师生角色等的影响, 提出了对中国数字人文教育发展的全方位建议。目前, 各大高校积极应对信息技术带来的教育体系、科研合作模式等多方面的变化, 在学科设置更新和科研方法磨合方面进行了颇多尝试。中国人民大学等高校在信息资源管理一级学科下设置数字人文专业或数字人文方向, 建设了一批数字人文研究中心、研究院或实验室, 极大地推动了数字人文领域的快速发展。武汉大学于2023年11月17日发布《数智教育白皮书》^[2], 率先提出体系化数智人才培养方案, 按下“教育数智化转型”的启动键, 以培养具备数字思维、数字素养和智算技能的数智人才。武汉大学副校长周叶中教授代表学校宣布, “为支持数字人文研究的可持续发展, 武汉大学在信息管理学院开设‘数字文化’本硕博贯通制实验班, 借此探索跨学科专业建设经验, 并为数字中国和文化产业数字化战略培养国家急需人才”。

武汉大学简帛研究中心主任、人文社科资深教授陈伟的《简牍整理研究中的数字技术支持》主旨报告, 对简牍整理研究中现阶段的实际研究问题、产出期望、技术诉求和面临的挑战等进行了深入的阐述。报告从简牍文献概念的辨析、数字化过程中的图像获取、简牍编连缀合和文字释读等四个方面展开, 通过具体实例, 展

示了简牍整理、图像获取、文本复原等一系列工作流程及技术路径。其中着重介绍了简牍数字化辨识的一些代表性成果和简牍变形、字迹模糊误识、异体字识别等尚待解决的问题,并特别展示了利用红外相机拍照技术进行简牍数字化的阶段性突破成果。

武汉大学信息管理学院副院长、文化遗产智能计算实验室主任王晓光教授结合科学领域内四种科研范式演化过程中数字信息技术深度参与的情景,以武汉大学文化遗产智能计算实验室的实践为例,从学术研究范式的角度对数字人文作为系统性范式变革的重要因素给科学研究范式革命提出的挑战和带来的机遇等进行了深度阐述。王晓光认为,“数字人文实验室作为新的科研组织机构和教学场域,在学术技术设施设备,构建、推动和传播新的学术研究范式,探寻新的学术知识传播方式等诸多方面的变化和转型中不可替代的功能和作用。”作为实验室主持人,王晓光的报告着眼于元问题的研究,无论是对于建设产学研一体化的学术体制机制,还是对于多领域、多学科协同构建新文科的实践,都具有理论意义与实践意义的启示。

1.2 AI 时代的智慧数据、文化科技创新

在数字技术日新月异的 AI 时代,信息与数据每分每秒都处于持续大规模增长状态,数据可信性成为一个不可回避的问题。数据可信问题包含数据准确性、一致性、完整性、可解释、可溯源、可使用等多方面,直接影响科学研究的可靠性、实践流程的稳定性,利用数据挖掘分析结果进行决策的准确性和适用性,以及受众信任的建立。来自美国肯特州立大学的曾蕾教授在《AI 时代可信智慧数据的生成——可能性及其挑战》的报告中,分享了她在人工智能背景下从事文化遗产和博物馆研究项目时面临的复杂挑战、全新机遇及相关心得。报告阐述了大数据向智慧数据转化的意义以及在数据结构化过程中创建机器可理解、可处理、准确可复用数据的必要性,着重强调了在以 ChatGPT 为代表的大语言模型浪潮冲击下多元数字智能技术并行不悖的重要性。整个报告以案例呈现方式展开,串联了多类型的智慧数据平台和软件,涉及文字处理、图像处理等技术运用,涵盖众多物质文化遗产和非物质文化遗产研究项目的成果。报告展示了在某种适当的情境下,项目所选择的技术平台、工具软件等在实现机器可读的数据资源和需要转换为机器可读数据的非机器可读原件预期的可能性空间。

文化与信息技术的双向融合催生了新的文化业态,与此同时,产学研关系也面临新的调整。北京邮电大学人工智能学院特聘研究员、交互技术与体验系统文化和旅游部重点实验室主任张树武教授,呼应新时代大变革背景下国家“文化数字化”战略部署和建设“社会主义文化强国”的决策,在《数字经济场景下的文化科技创新》报告中结合中国国情,阐述了文化科技创新的新内涵、新需求,以及当前数字经济的新场景。他从宏观战略、组织方式、产业机构及发展理念等层面,重申了文化数字化作为重要突破点的重要性和任务。武汉大学人文社科资深教授马费成也在致辞中报告了武汉大学大数据研究院集数据科学研究、数据平台建设与服务为一体,重点聚焦金融、商务、人文、健康、社会治理与传媒等研究领域,致力于创建国家大数据战略研究与应用开发领域创新平台的情况。

上海大学党委副书记、纪委书记段勇教授的《GLAMS 点亮元宇宙的星空》报告从“GLAMS”概念的发展与拓展出发,阐述了从早期侧重图书馆(L)到美术馆、图书馆、档案馆、博物馆(GLAM)等多个领域协同研究,并随着科学中心(S)的深度融入,各类文化机构“科技化”与科研机构“人文化”的双向融纳趋势,而该趋势即将、正在和必然改变未来文化传承、传播和焕新的生态。报告认为,随着数字技术和人工智能的发展,人类文明正

处在未来的拐点,科研、生产以及社会生活的各个领域正在实践逐步构建一个与现实世界融合的虚拟世界,这将对人类的文化传播和文明传承产生至关重要的影响,甚至可能左右未来人类的进化和社会的进步。

2 分论坛概述

本届数字人文年会共十个分论坛,也即十个议题,收到摘要投稿共 297 篇。经年会学术审查小组两次筛选,最终选定 83 篇,论文作者向大会提交了完整论文。各分论坛投稿占比见图 1。

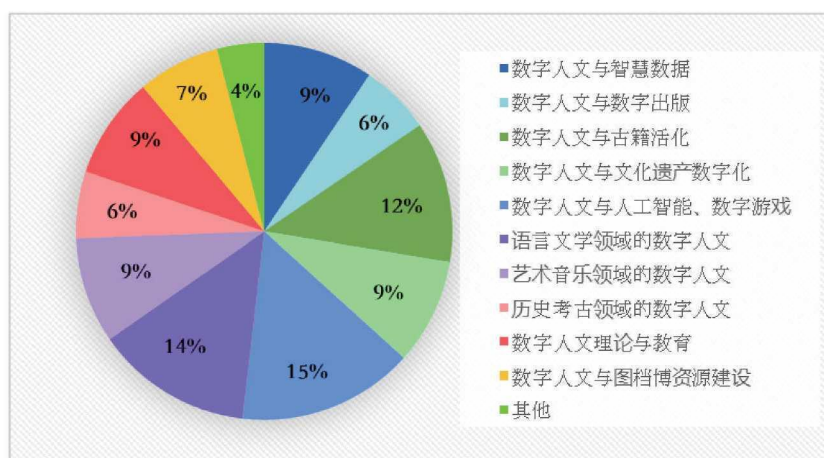


图1 各分论坛摘要投稿分布比例

2.1 数字人文与智慧数据

在当前国内,不论是高校研究机构、图书馆、档案馆、博物馆,还是企业或社会机构,大多以数据库和知识库的建设为切入点开展数字人文研究。虽然并非所有研究者都明确使用了“智慧数据”这一概念,但在实践中已基本上默然采用。所谓智慧数据,指的并非是数据本身具有智慧属性,而在于数据中蕴含能够被提炼的知识、应用智慧等。换言之,智慧数据意味着能够从任何规模的此类数据中获得重大的洞察力。相比于具有规模(Volume)、速度(Velocity)、多样性(Variety)、易变性(Variability)和真实性(Veracity)这5V属性的大数据,智慧数据具有可信(trusted)、情景化(contextualized)、相关切题(relevant)、可认知(cognitive)、可预测(predictive)和可消费(consumable)六大特性。^[3]

从本次年会“数字人文与智慧数据”分论坛的论文投稿内容来看,目前国内数字人文智慧数据的研究与尝试主要集中在知识抽取与知识对齐,知识关联与组织,知识融合与呈现以及知识服务四个方面。这四个方面相互交织,共同构成了目前数字人文领域智慧数据开发与建构方向的研究框架,所涉及的方法和技术多采用“国际通用标准+特定针对性工具”的形式。

知识抽取与知识对齐是当前文化遗产相关知识库构建的基础。通过运用国际通用标准和特定针对性工具,研究者们致力于从各种文本和数据源中提取文化遗产的关键信息,并将其对齐到统一的数据模型中。这有助于建立起一个更为一致和可比的知識基础。中山大学廖嘉琦的《基于大语言模型对话的藏书题跋信息抽取与知识建构——以黄丕烈藏书题跋为例》以藏书家黄丕烈所撰写的藏书题跋为研究对象,与大语言模型工

具展开多轮对话,对藏书题跋数据进行预处理、信息抽取以及知识建构,初步构建了藏书题跋知识本体模型。中国人民大学盛一涵和谭可婕的《面向制度史的中国古代职官本体构建探索》以历时和共时两种视角,在官制辞典和史料的记载范式及国内外前沿本体构建方法的基础上引入 RDF 四元组,尝试实现复杂职官制度的统一理解与表达,摸索职官本体、历史本体、制度本体的多层次实现。

知识融合与关联、组织与呈现能够辅助知识逻辑化,促进隐藏知识的发现。知识关联与组织重视通过建立知识之间的关联关系来尝试构建相对更为完整和多维的文化遗产知识网络,满足不同层级的知识检索、访问、使用等需求。相较于知识关联与组织,知识融合与呈现多了一层创新的思想,强调将多源、多样的知识进行融合,并以直观、易理解、新颖高效的方式呈现给用户,但就技术手段与研究方法方面而言,二者多有重叠。南京大学数据智能与交叉创新实验室研究团队的《基于知识重构的史料资源数据线索发现研究》结合人文研究方法与数字人文技术体系,提出基于数据线索的研究框架,并以《拉贝日记》为例重构史料资源所含知识、挖掘数据线索,论证数字技术如何介入人文研究及其重要性与可能性。广西大学唐振贵、武汉大学罗锦坤的《星空记忆:中国古代天象记录智慧数据构建框架研究》综合使用大语言模型技术与知识图谱方法,构建由史料层、注解层、关联层、融合层和应用层组成的古代天象记录智慧数据五层框架,探索提升天象记录数据形态的智能知识组织框架。两项研究显示出数字人文方法在构建古代科技文献和近现代史料文本智慧数据过程中的可行性。

华中师范大学周树斌等的《四川清音知识表示及语义关联构建研究》针对四川清音基本知识结构进行本体建模,结合四川清音现有数字资源,从实证层面探索实现领域知识的语义关联和细粒度解构。华中师范大学吴艳飞等的《中国古代纹饰知识图谱构建与知识发现研究》,对楚国动物纹饰资源的知识内涵进行知识抽取、融合、关联、重组,构建了领域知识图谱。研究者的成果为文化遗产领域其他资源的数字人文研究提供了参考借鉴,也证明了知识融合与关联、组织与呈现可为文化遗产领域中多类型、多领域资源的细粒度、全方位多维组织与知识发现,提供新的方法和思路。

知识服务是将研究成果转化为实际的服务,是从知识抽取到呈现的完整过程的最终目的,集前三个层面之大成。如数字图书馆、文化遗产数据库提供研究工具、空间定位、结果呈现等,以为学者、文化工作者和公众提供更为便捷和可访问的资源,促进现有资源可检索、可访问、可相互操作以及重复使用。金陵科技学院葛怀东等在《数字人文视域下家谱数字资源的开发与建设》分享中提及家谱资源藏与用的矛盾,表示数字人文方法为家谱数字资源的开发与利用提供了新的途径,同时指出针对新条件与机遇,家谱数字资源开发须有适配、灵活的新策略。

总体上看,该分论坛分享的 8 篇论文,虽然作者来自不同的区域和单位/机构,研究对象也各不相同,但开发理念、标准、选用平台、技术路径和方法,基本上是同构的。上述四个层面的研究虽然取得了显著进展,但仍然存在一些明显的不足,以不同领域的知识融合问题最甚,尤其是在跨学科合作和数据整合方面。不同学科有各自的研究方法和理论框架,这意味着人文学者与技术研究者展开同一个或者同一类数字人文研究课题时,双方的切入点、关注点以及思维模式天然存在着很大壁垒。^[4]因此,跨学科融合要求研究者不仅熟悉自己专业领域的方法,还要能够理解和整合其他学科的研究方法,形成更为综合和全面的研究方法。虽然这一点共识已经达成,但实践所依赖很多的因素尚待进一步完善,包括研究者的知识结构、体制机制、工作条件,等等。

2.2 数字人文与数字出版

数字出版和数字叙事是数字化发展浪潮的重要产物,也是推动变革的要素,数字人文方法的融入为二者带来了新的可能性。从本次年会“数字人文与数字出版”分论坛投稿情况来看,数字人文工作者们现阶段更关注于数字技术参与的知识呈现和展陈的新方式、所带来的新效果,以及这一过程所需的数字装置相关理论与技术、新的数据架构需求之于受众全新知识获取途径的颠覆性意义等,同时也意识到了数字藏品发展中存在的诸多问题和隐患。

数字叙事是信息化与数字化催生的产物,数字装置艺术增加了数字叙事的可能性和多样性。20世纪60年代和70年代文学理论的变革、跨学科研究的兴起、文化相对主义思潮的盛行,以及数字媒介与技术的发展等多重因素共同作用,催生了20世纪后半叶逐渐形成的新叙事学。新叙事学强调跨学科的研究方法,涉及文学、电影、社会学、心理学等多个领域,试图超越传统叙事学的限制,更全面地理解和解释不同社会文化背景下的叙事现象。在信息化与数字化驱动下,数字叙事成为新叙事学一个新的研究领域,它利用数字技术和互联网工具来创造、传播和体验故事,与传统叙事手段和手法相比,具有多媒体融合、非线性结构、交互性强、体验感强、社交分享等特性。在当下的AI时代,数字叙事实际上是数据驱动叙事,可利用大数据和人工智能技术,根据知识传递需求或观众的个性化需求和行为数据进行定制,提供更丰富多元、更符合用户喜好的叙事体验。

数字叙事为传统展览、阅读方式和文化遗产的传承带来了全新的可能性,剖析其中产生的新文化艺术现象和受众的身份、体验转变成成为数字人文的重要研究内容。中国科学技术协会国家科技传播中心陈蕾在分享《数字装置艺术在沉浸式展陈中的叙事研究》时认为,物的性状改变促使沉浸式虚拟展览的意境选择和叙事结构都发生变化,身处数字装置艺术体验中的受众不仅是信息接收者,同时也是叙事者。北京印刷学院郑春晖的《从接触到具身:浅谈阅读的数字叙事体验》以身体、交互和叙事语境三者关系为切入点,研究数字时代新型阅读方式下读者具身参与和场景构建的演变和特征。

数字人文方法和技术为数字出版带来新的生机和挑战。山西大学肖珑、乔晓玉的《数字人文方法在融合出版中的应用——以五台山壁画出版平台为例》选择充分体现中西文化交融时期的五台山壁画为研究对象,在“都柏林核心元数据”和国家文物元数据总则的统一标准下设计五台山壁画元数据架构、对图像进行处理,推动五台山壁画出版平台的搭建,实现五台山壁画的知识可检索。北京印刷学院张孜晖在其《“出版+数字藏品”发行研究》中指出,数字出版领域当下面临版权模糊、内容质量参差不齐和脱离实体经济等问题,结合多感官协同、区块链等新技术的实践及更加深入的研究方法,有望推动数字出版领域更为全面的发展。

就该分论坛上分享的研究成果来看,叙事学、图像学、解构主义、融合出版等本来风马牛关系的理论,在探讨数字人文视域下的数字出版和数字叙事时候却呈现出合和通变的灵动性,这体现了该类研究问题具有的高度综合性和纵深感。

2.3 数字人文与古籍活化

古籍活化工作已经上升为一项国策。2015年习近平总书记提出“让书写在古籍里的文字活起来”;2022年4月中共中央办公厅、国务院办公厅印发的《关于推进新时代古籍工作的意见》明确提出要完善古籍工作体系、提升古籍工作质量。^[5]相关方面的参与者也为年会提供了可圈可点的案例。

古籍活化研究对象类型广泛且学科多样。此次“数字人文与古籍活化”分论坛投稿论文的研究对象类型包括传世典籍、出土文献、纸本、碑帖、金石,内容涉及农业、中医、文学、造字等,充分展示了数字技术辅助的古籍活化工作在不用学科中的相对通用性。西北大学周知等的《基于残差网络的碑帖年代识别方法》运用图像增强、卷积神经网络、深度学习等技术,针对明、清、民国三个时代的碑帖图片进行识别与内容挖掘。南京农业大学吴梦成等的《面向古代农书的翻译与知识图谱构建研究》利用知识标注、实体和关系抽取等技术所抽取古农书中的知识,构建领域知识图谱,探索古农书的翻译与理解方式。中医类古籍资源库因其学科的独特性质和内容而特别引人注目,南京大学张君冬与郑州大学杨松桦等的《AIGC 赋能中医古籍活化研究:Huang-Di 大语言模型的构建及应用》借助已有开源模型构建了中医古籍生成式对话大语言模型。从论文投稿情况来看,中医药高校在数字人文领域较之前更为活跃,综合类高校及研究所对中医类古籍关注度也逐步提升。

总体而言,本次年会呈现的研究成果显示长期以来以古籍文本内容为研究重心的态势逐渐被打破。

首先,关注点逐渐分散,研究内容日益丰富,扩展至古籍活化体系的不同层面,并带来新启示。西北大学刘睿等的《出土文献集外字造字法实践路径研究》着眼于集外字的特殊性,从“造字法”的角度切入,综合利用 Unicode 字符集、OpenCV 库、Potrace 算法工具和 FontCreator 软件等,尝试为识别处理集外字造成的挑战寻找出路。北京大学王凤翔等的《古籍目录的量化研究框架》表明数字人文不仅在知识产出方面具有生产力,而且能够为知识产出的条件、路径及过程提供适配的方案。南京农业大学孙悦帆等的《我国社会力量参与古籍保护工作的路径选择》显示古籍保护工作已不再是学术研究机构 and 学者的专利,业界、社区和公众均成为参与主体,古籍活化的社会公众性和服务功能正在切实实现。

其次,研究方法与技术选择日益多元。比如 OCR、自然语言处理、自训练模型、图像分类、卷积神经网络、大语言模型等技术已经被古籍研究领域数字人文工作者广泛使用,研究者们通常会在已开源模型的基础上配合针对性的语言模型,以提高建构模型的准确率。山东理工大学范颜铄、中国人民大学周晓英等以《聊斋志异·司文郎》为实操案例,利用大语言模型和关联数据等技术与方法对该类古籍资源的关联数据发布过程展开研究,在文学类古籍文本知识的链接、智能应用以及知识可视化呈现等多方面进行了系统性预设与生成的尝试。南京农业大学齐月等则使用词汇统计、语言计算等方法,结合大语言模型文本表示方法和词向量等技术手段研究《春秋》三传语体风格的关联与差异。

2.4 数字人文与文化遗产数字化

本次年会“数字人文与文化遗产数字化”分论坛提交的论文关注点多集中在文化遗产保护如何面向中华民族历史溯源与接续研究,数字化手段如何为文化遗产保护提供更多元的选择和全新的机遇。目前该领域,尤其在延长文化遗产生命周期和跨越历史时间与时代同频方面,已经开始由理论架构进入具体项目的实践^[6]。不过,现阶段文化遗产数字化的相关工作流程表现出高度的同构性和同一性。这可以从两个方面来理解,一方面,不同地区、不同机构、承担不同职能的研究单位选择侧面不同的项目组,在各自为战的情况下形成了日益统一的工作规范,为未来统一的数字人文大平台建设奠定了比较好的基础;另一方面,高度同质化的工作可能意味着创新所需要的多元化工作生态相对缺失。浙江大学高荧、武汉大学姜钰璇、西安建筑科技大学何冲、南京理工大学孙绍丹等分别从文物复原、遗址景观重现、文化遗产语义增强等方面探讨现阶段研究与实践中存在的共性问题。中国人民大学陈晓婷围绕《我国文物出境展览的时空分布特征及其成因》,向与会者分

享了我国文物出境展览的时空分布特征及背后的政治环境、社会经济等多方面影响因素。

基于进行中的文化遗产数字人文研究或项目的文化再生产课题也已经提上日程。西北大学李姗姗等的《面向实数共生的关中拴马桩文化遗产 3D 建档与开发》对团队关中拴马桩 3D 建档与相关领域知识聚合的工作进行了阶段性成果分享。华中师范大学翟姗姗等利用知识图谱对国内外饮食类非遗研究展开规律性挖掘。广东工业大学陈晨等的《设计赋能遗产活化》以图像叙事学为理论基础建立图像符号数据库,探寻数字化手段下佛山西樵世界灌溉工程遗产的价值阐释与展示路径。

以上研究也反映出从宏观来看我国文化遗产数字化工作仍处于起步阶段的状况,数据资源共享、数字化采集工作、数字化展陈等各环节存在着严重的资源孤岛问题。缺乏统一规划和科学顶层设计、侧面优先的单一项目各自为战,成本高且难度大、表现力不强、交易机制不完善等已经成为该领域数字人文工作不得不面对的问题,如不及时克服,势必严重限制文化遗产数字化和文化科技创新的纵深融合^[7]。

2.5 数字人文与人工智能、数字游戏

随着人工智能的长足进步,关注人工智能应用的研究数量增长显著,这在本次年会上也有充分显示。

大语言模型在人文社科、教育等领域得到广泛应用,机遇与挑战并存。南京大学刘江峰等的《AI 赋能计算史学研究:Thinkers 思想家大语言模型的构建及应用》以百川大模型作为基准模型,针对《思想家评传》建立思想史领域大模型。武汉大学杨思洛等则利用主题模型,预测元宇宙热点与前沿。武汉大学杨逸云等在《ChatGPT 下的学术教育:国外应对策略与启示》的分享中分析面对生成式人工智能工具井喷式发展国外学术教育的应对策略,总结了新技术推动的教学模式、师生关系等方面的变化与特点,并思考国内教育应对之策。中国艺术研究院彭明君与清华大学卢睿的《虚假共现:AIGC 偏见探源及其与人类偏见的交叉反思》综合运用机器学习、认知科学和语言哲学的方法,深入探究“偏见”生成的根源,强调 AIGC 偏见是人类认知及认知偏见生成的镜像,为 AIGC 影响下的偏见治理提供了新思路。

生成式人工智能也逐渐融入电子游戏设计当中。都柏林圣三一学院唐慕琳的《“制造玩家”——从数字人文视角对比 MOBA 与 MMORPG 游戏类型 AI 应用模式下的游戏文化与玩家塑造》针对人工智能带来的新旧机制变化,对玩家和游戏文化的重塑展开讨论。历史类电子游戏被视作“弥合数字史学与主流史学张力的突破口”——中共江西省委党校卢雅怀在《历史游戏研究作为数字史学发展新方向:一份研究综述》中梳理了电子游戏中的历史叙事及其对历史研究的影响,认为“历史游戏研究构成了历史教学科研面对数字时代的重要方式”,是“重要教学工具”和“学院与公众沟通必要途径”,更是“做历史研究的新工具”。

人工智能生成内容在文化创意产业中应用的典型之一,是博物馆与游戏公司间的合作,包括利用 AIGC 进行博物馆游戏设计,通过 AI 对藏品数据进行分析 and 处理,以及通过 AIGC 打破博物馆开放 API 元数据的壁垒等。莱斯特大学蒋涵等的《AIGC 支持的协同设计:基于博物馆数字化藏品的游戏机制》探索了在打破博物馆开放 API 元数据间壁垒的同时提升叙事线索可及性的方法,尝试构建博物馆藏品 API、大型语言模型和游戏开发者、设计参与者、利益相关者皆关联在内的开放式技术网络,希望通过改变叙事设计促进游戏中玩家的个性化探索与交互体验。

元宇宙依旧是研究热点。武汉大学杨思洛等在《基于 LDA 主题模型的元宇宙热点与前沿探测》中强调,虚拟世界与现实世界的跨越联通、虚拟现实和深度学习的深入应用等方向是未来的重要趋势,主要涵盖区块

链、深度学习、物联网、虚拟现实和显示器等技术。

2.6 语言文学领域的数字人文

数字技术的自动化和智能化推动了语言文学领域数字人文工作的快速掘进,主要表现在研究材料的多元化和解释框架的开放性。本次年会“语言文学领域的数字人文”分论坛收到的论文不仅数量第一,平均质量也更胜一筹。这些丰富多彩的研究覆盖了文学、历史、地理、艺术等多个学科,且不乏多学科交叉与融合,研究材料包括中国近现代文学、古代诗歌、古典戏曲、近代报刊、地域童谣等。

针对文本特征的研究越发具体,越具体则意味着越深入。清华大学陆泉宇在《“上尾”“鹤膝”与近体诗平韵传统的形成——基于统计方法的考察》中对近体诗的韵律特征进行深入考察,通过四声标引和统计分析探讨永明体到近体发展中“平韵”传统的形成原因和图景。同为规律性研究,清华大学郝若辰等的《说著同光已惘然:基于预训练模型的晚清吟坛风格流别》通过运用预训练模型、词向量、聚类等方法对同光体的风格区分度进行量化考察,希望揭示同光体实质上存在复杂的意义层级与风格指向,验证了同光体在风格上与其他宋诗流派的差异,并强调了历史认定上的问题。南京师范大学林子炀等通过对《万国公报》中高频词和特色词的计量分析,挖掘报刊内容的主题演变及此前被忽视的关键信息。也有研究者关注到此类研究中研究者角色的丰富与研究思维方式的转变,如武汉大学张亚静在古典戏曲研究平台搭建过程中受到启发,在《艺术家与工程师》一文中提出以文学为中介的古典戏曲数字人文研究需要“艺术家”与“工程师”的双重思维。

从语言文学延伸至历史和文化研究的过程中,研究者们多聚焦于网络关系与时空分布。华中师范大学繁爽等在《数字人文视域下古代诗人时空分布与多维知识发现研究》中从地理角度分析诗人文学创作及其情感表达的内在关联。闽南师范大学吴泓润等的《数字人文视野下闽南童谣的网络分析》对闽南当地特色童谣的整体特征、度分布、中心性和社区检测进行可视化和定量分析,探究童谣词汇使用特征及其与宗族文化、传统习俗等之间的关联。密歇根大学佟遥在《从零到一,女性团体在公共领域和政治空间的初次涌现——戊戌维新到民国初年(1875—1913)的女性团体 GIS 呈现与社会网络分析》中结合地理信息系统与社会网络分析方法,以《辛亥革命时期的妇女团体》和《中国妇女运动百年简史》为主要数据抽取来源,探究晚清民初女性团体的发展及其在公共领域和政治空间中的分布特征。

2.7 艺术音乐领域的数字人文

“艺术音乐领域的数字人文”分论坛投稿论文内容涉及电影、绘画、音乐、书法等艺术类型,通用技术包括知识图谱、聚类算法、自然语言模型、社会网络、本体技术、数字孪生等。哈尔滨工业大学(深圳)陶婷等通过色彩计算方法分析电影画面内容、色彩元素等,深入挖掘电影《大红灯笼高高挂》中蕴含的多维信息。来自跨烟台科技学院、韩国江原大学、扬州大学、曲阜师范大学、马来西亚理工大学等高校团队的随成龙等的《数字人文视域下雅集绘画作品的知识图谱构建研究》以雅集为研究题材进行知识图谱构建,以期实现集领域知识的组织、语义链接、检索和可视化展示为一体的用户服务。中南财经政法大学石锴文等研究古琴减字指示的乐曲节奏规律,并尝试实现自动化与智能化的记谱方法。西北大学李姗姗等的《面向数实相生:秦腔非遗传承人图谱与口述档案建构》以“封存秦腔传承人的活态记忆与隐性知识”和“活化秦腔艺术”为目的,设计了针对性口述档案元数据方案。研究者们尝试为各种存在样态的传统文化的数字化保护和传承开创新途径。这些案例

也突显了数字人文艺术研究中跨学科合作的重要性。而在数据质量、标准化、技术更新等问题上,构建专题性数据库、采用元数据方案等是大多数研究者认可的有效途径。

2.8 历史考古领域的数字人文

“历史考古领域的数字人文”是分论坛的名称,其涵盖范围较常见的“数字考古”更广,内容更丰富。“数字考古”已经被学界和业界确认为未来中国历史文化遗产保护的新途径,^[8]个别院校可能已经将之列为课程。可大胆预言“数字考古”将进入相关专业目录,不排除将来成为独立学科,而非仅仅是作为纯粹工具在考古中使用。本次分论坛的论文研究对象包括出土器物、服饰文化、建筑工程管理、宗教传播以及军事防御体系等,二重证据法、社会网络分析、聚类分析是研究者们目前最为偏爱的三种方法,ArcGIS、3D 建模等技术手段是研究古代文化器物和历史遗迹过程中常用的辅助工具。

武汉纺织大学董冰卿的《虚拟仿真视域下马山一号楚墓彩绘着衣女木俑服饰复原》运用二重证据法,对马山一号楚墓彩绘木俑服饰的纹样色彩、形制、面料进行分析,利用2D结构复原、3D建模等技术复原服装色彩,以此深入挖掘楚国礼制僭越的历史信息。湖北大学吴志勇等的《基于地理空间的古代印度摩竭陀国佛教传播可视化研究》以ArcGIS为平台,在可视化空间探寻摩竭陀国佛教传播活动,从传播学与文化地理的角度对佛教传播展开了研讨。为应对人文社科领域对文本数据和地理空间栅矢数据等进行融合性语义分析的强烈需求,西安地图出版社白江涛在《多源语义下的历史地名时空辨识与信息化建设——以黄河历史河务地名数据库建设为例》中提出将文本语义转化为具有时空载赋的矢量信息的思考和多源地理空间数据关联方法。天津大学陈清越等针对《随工日记》抽取数据并进行社会网络分析,试图揭示清代陵寝工程管理背后的人际关系和复杂信息决策模式。天津大学卞俊杰等的《基于军事地理学的明长城真保镇军事聚落空间防御力评价》以内边长城真保镇及其管辖范围为研究对象,以主观和客观分析法、空间聚类等方法分析各路段防御体系的效果。

无论是服饰器具还是遗址古迹,都是文化的载体,其残缺破损、历经风霜的物质实体蕴藏着礼乐文化、社会风俗、地域风尚、交通运输、地理气候等丰富的历史信息。数字技术与学科传统研究方法相结合的研究路径提供了考察过去文明的新视角,有助于重建更为完整和立体的历史图景,形成关于过去文明与文化更全面、深刻的认知,同时也为相关领域的保护、传承和研究提供了丰富的实验价值。

2.9 数字人文理论与教育

数字人文理论研究在当前的必然性、紧迫性,不仅在于理论是实践发展到一定阶段之后必须的经验总结和继续研究必需的科学预设,还在其基础性、原生性、独特性是对特定事件或现象发生的基本原理和原因进行挖掘的无可回避的选择。本次年会“数字人文理论与教育”分论坛的论文主要涉及新型阅读范式,国内外数字人文领域研究热点及主题演化趋势,数字人文学术共同体的建设与困境,数字人文研究中数据可持续访问与重用问题四个大的方面。

上海大学图书馆夏明宇借《细读、粗读与远读:从传统阅读到数字人文的范式演变》分享了三种阅读范式在数字人文语境下的关系与内涵,试图揭示阅读范式生成与演变的内在逻辑。南京大学张逸勤团队分享的《国家战略情境下我国数字人文学科建设:分析与展望》和上海大学骆雨辰团队分享的《智能时代的转变:近十年数字人文领域主题热点和演化趋势——基于国际数字人文大会的分析》,使用KeyBERT关键词抽取工具和

词频分析、共词分析、聚类分析方法分别总结了国内和国际数字人文研究主题的演变,预判发展趋势。学界、业界和有关机构探讨智能技术在知识挖掘与分析、文学史学档案可视化、LAM 合作与众包协作模式等方面的赋能作用,数字人文研究日益呈现出关联、开放与融合特征。

数字人文研究跨学科、跨领域、跨圈层的特点意味着数字人文学术共同体建设的必要性。中国人民大学张旭在《数字人文学术共同体建设的现实困境与实践进路》中提到主体、场域、活动、空间、机制皆为影响共同体建设的基础要素。正视当前困境,探寻数字人文学科学理化、专门化、国际化、本土化道路是一个迫切而又事关长远的议题。南京大学练靖雯等的《枯木逢春犹再发:数字人文项目数据可持续访问和重用的案例研究》通过意大利 RESTORE 项目、牛津大学“Digital Scholarship”平台、伦敦国王学院的 King’s Digital Lab 以及剑桥大学的 Cambridge Digital Humanities Lab 四大典型案例,调研分析各国数字人文领域公共数据基础设施建设与各方资源分配现状,以期为中国数字人文建设可持续访问和可重用的基础设施平台等找到明路。

在数字人文教育方面,北京师范大学张宁的《面向大学生传统文化教育的 VR 古籍活化设计体验效果比较研究》分享北京师范大学 VR 古籍设计的实践成果以及在赋能传统文化教育方面的不足与进一步的优化方案。湘潭大学肖可以等以中国六所高校为例调研数字人文教育的现状,分析了课程类型、高校图书馆支持、教育服务体系三方面的情况,并提出了改进建议。英国伦敦大学学院高瑾以该校数字人文教学中生成式人工智能应用的情况以及存在的问题为例,探讨了数字人文教学中融入生成式人工智能的策略与准则。天津师范大学黄晓霞的《数智时代古籍数字人文的人才培养与课程建设》从古籍智能处理、古籍智能修复、古籍知识挖掘与古籍数字活化四个维度构建了古籍数字人文的课程体系,并结合天津师范大学古籍保护研究院开设古籍数字人文课程的探索实践案例,提出了灵活适配的教学建议。综合来看,当前我国数字人文教育在国家政策引导下逐步发展,对其不仅需要关注战略性指导和理论建设,还必须着眼于实践研究,特别是实际应用和用户体验,这需要学界、业界,甚至金融界的多方面合作携进。

2.10 数字人文与图档博资源建设

与高校研究导向的项目和课题不同,图档博机构的性质决定了其研究最终落脚于面向社会公众的知识服务与知识传播,以普及文化、历史、艺术等领域的知识,满足公众的了解和学习需求。故宫博物院的《基于 OAIS 模型的“在线文物数据库”系统框架设计与实现》、上海图书馆的《名人文献的数字人文场景构建研究——以宋庆龄书信为例》、贵州省毕节市档案馆的《数字人文发展视域下贵州红色档案资源保护与开发利用策略研究》及重庆大学图书馆的《高校数字记忆的建设路径——以重庆大学为例》,是对各自馆藏资源进行数字化存档的实践总结,研究者们力求在有限条件内最大程度实现资源可查找、可访问、可互操、重复使用;通过挖掘自身馆藏的特色,建立特色资源库与元数据方案,强调独特的文化、历史或艺术元素;还通过虚拟展览、人机交互等方式实现在线展示与互动。即同时进行馆藏资源保护与开发利用的研究与实践,这一努力服务大众的事业在数智化时代必将获得更广阔的发展空间。

与此同时,关于图档博机构如何依靠自身优势,通过特定项目整合高校相关专业人员、获取企业技术资源和金融等各方面的支持,进行着眼于更长远考量的体制/机制建设方面等问题的研究和建设略显不足,而这恰恰是推动国家文化数字化战略在图档博领域真正落实不能不正视的问题。

3 年会项目总结与分析

本届年会项目的征集采取专家邀请和团队/个人自行报名相结合的形式,以期收获关注重点又不遗漏明珠之效。从最终结果来看,预期基本达成。总共征集到 35 个数字人文项目,其中团队项目 33 个,个人项目 2 个,最终入选项目 29 个。项目评选设置最佳项目奖 4 个,并从平台工具性与可适性、题材新颖性与价值性、呈现设计创造性与适用性、内含技术先进性与精确度、项目体验感,以及人机交互设计互动方式六个维度进行考量,分别设置最佳平台奖、最佳题材奖、最佳设计奖、最佳技术奖、最佳体验奖和最佳互动奖各一,优秀题材奖、优秀设计奖、优秀技术奖、优秀体验奖、优秀互动奖各三,以及优秀平台奖四个。

本次年会所征集项目总体数量和前几届相比虽然有所下降,但质量和深度融合方面越来越向精准化、精品化、优质化发展,具体表现为平台和具体交互设计等可操作性显著增强,项目实践预期更加明朗,专业性更强等。

3.1 团队地域及机构分布

除 1 个海外团队外,投稿项目团队分布于国内 15 个城市,北京、上海、武汉、广州四大城市的团队贡献了 35 个项目中的三分之二。投稿项目团队来自 34 个机构,包括 26 个高校科研院所或实验室、8 家业界机构以及 1 个人工作室。值得一提的是,投稿中有 7 个跨校或跨机构合作项目,以及 5 个校内跨实验室、研究所或跨学科合作项目,在跨学科交流与合作方面较上几届有显著提升,再次说明数字人文激发了不同地区、不同机构协同作业的意愿与合作动力,为研究者及参与者提供了开展全方位深度交流的条件。这一趋势也大大促进了产学研相结合的新型科研模式由理念到实践的转变和升华,初步奠定了众多高校、业界机构以及个人工作室在共同探索数字人文领域时良好的合作互惠基础。这一趋势预示着未来数字人文领域的发展将更加多元且具有协同性。

3.2 类型区分

从项目类型上看,现阶段数字人文项目以资源聚合及平台展示、工具性平台、数字游戏与立体展陈为主。在会议录用的 29 个项目中,此三类分别占比 51.7%、31%、17.3%,这一统计一定程度上体现了现阶段国内数字人文的研究需求和关注重点。资源库、数据库建设仍不完善,海量、多模态的有价值数据仍有待挖掘并结构化,因此资源库建设通常着重强调数据的整合与展示,致力于为学术研究提供更广泛、更丰富的素材。工具性平台的优化虽然一定程度上能够反映信息技术融入人文学科研究的结合部与切入点,但是现阶段的工具性平台在精确度、涵盖范围广度等方面,若想满足人文学科及社会科学研究需求仍要进一步努力。如果说前两类项目更倾向于服务专业研究人员或爱好者,重视人机交互设计的数字游戏与立体展陈则更面向社会大众,目的在于传播文化内容。本届年会此类项目虽然仍占少数,但与前几届同项目内容相比,已然体现出在互动方式、创意、趣味性传播质量上的提升和关注度逐年增长的态势,凸显了国内数字人文研究视野的逐渐开阔,即不再局限于科研目标,而开始考虑技术与人文深度融合过程中文化的社会传播。这一发展趋势为数字人文领域注入了更多创新活力,也反映了数字人文日益增强的社会影响力。

3.3 资源来源及研究对象

数字人文项目类型及资源来源很大程度决定了其研究对象。本次年会所征集的项目数据来源多以古籍、

文献为主、图像为辅,纯图像资源辅以文本数据者占比相对较少。资源库建设团队在获取数据时多遵从可访问、可获取及项目团队具有较高使用开发权限、数据相对熟悉且完整全面的原则,因此数据资源一般来源于项目团队所属单位或合作单位,呈现出单位性质和地域特征,开放性和共享性受到限制。比如重庆大学图书馆与北京爱迪科森教育科技有限公司合作的“AI 助力高校历史的数字再现与智能挖掘——重庆大学数字记忆平台”以重庆大学各类相关资源为基础,对珍贵历史档案、文献资料、学术成果、图像影音、建筑文物等进行全面的数字化工作,从而形成一套与重庆大学相关的相对完整的数字记忆体系,以服务于该校历史与学术文化资源的数字化重现、长期保存、管理整序,并进一步支持知识再生产和学术交流活动。又如闽南师范大学物理与信息工程学院创建的“闽南文化的数字化模型及其示范应用”平台,该平台通过融合人工智能技术与典型闽南文化元素,辅以旧照片智能修复、风格迁移智能剪纸、表情合成等设计与功能,尝试实现操作方便且兼具个性化与趣味性、针对闽南文化数字化应用的简易人机交互平台。部分项目虽未被本次会议录用,但仍然具有同样特征。这也反映了国内数字人文经过十余年的发展,在关注点和研究侧重上正逐步走出粗放式的、经典文献扎堆的初级阶段。不同地区在选择项目时,开始侧重考虑本地区的政策支持、学术支撑、技术基础和地域文化特点等要素。正是各项目鲜明的地域特征与具象化、可操作、受支持、易深入、有特色和不可替代的特点,使数字人文得以形成燎原之势。

除基于机构优势和地域优势获取的数据,开放数据及已相对成型的资源型数据库也是现阶段资源聚合与平台展示类项目的主要数据来源。由北京大学智能学院团队开发的“大运河流域名人行迹可视化”网页即是综合利用开放数据及数据库的典例。该项目筛选、整合、重构了由哈佛大学开发的“中国历代人物传记数据库”(CBDB)与四川大学王兆鹏教授领衔开发的“唐宋文学编年地图”中大运河流域范围内的名人数据,将人物轨迹变迁与大运河流域变化相结合,利用可交互时间轴设计,实现以时空轨迹呈现千年运河背后的历史脉络;并且使用基于网页端的交互设计降低操作门槛,使用户能够根据对细节程度的不同需求总览全貌或查看单人轨迹视图,探索大运河流域的文化变迁与具体人物轨迹背后的故事。再如个人项目“中国绘画视觉探索”,该项目基于含 21344 幅中国艺术作品的“Artlib 世界艺术鉴赏库”中的网页浏览与分类查询工具,充分利用大语言模型进行辅助,实现了非线性交互数据的可视化。

3.4 成果开放性与传播

文化机构向社会大众进行文化传播的方式已越来越多元化和趣味化。随着技术的进步及思维方式的拓展,传播形式不再局限于传统实物、图像、文本、影像等的展陈,线上线下相结合、实体虚拟相结合已然成为趋势。国内数字人文相关基础设施建设和研究起步较早且资源、政策、人才、技术等均占优势的故宫博物院、上海图书馆、位居前列的美术学院,均是这方面实践的领先者。“数字文物库”平台是故宫博物院长期对外共享院藏文物数字资源的重要窗口,也是及时展示数字与人文研究成果的通道。项目团队在不断丰富文物资源数量和资源模态的技术路线基础上,增加语义增强策略,使“数字文物库”成为多模态、高质量的文物资源服务平台,在文物资源公开规模、影像品质、数据规范、知识检索等方面具有一定的行业代表性;所开发的“中国古代可移动文物知识图谱”为探索数据开放与交换奠定了基础,不断推进该平台从“资源服务”向“知识服务”的转型。

此外,诸多利用沉浸式展陈与创新性人机交互技术对特定知识进行传播的项目,已经具备了相当的市场意识,表现为从大众熟知的内容入手,从其已掌握的知识点切入,以新的形式进行新旧知识的关联,引导受众在提高理解的同时吸纳新知。上海图书馆投稿项目之一《“雪泥鸿爪”苏东坡 VR 游戏体验》是 2023 年 8 月该馆“灵境石语”展项中推出的一个体验性质的沉浸式 VR 游戏。该游戏基于对文献的解读,进行 VR 的可视化

图像设计,运用取意工笔山水和南宋文人墨客法帖中“碧空卷微云”“孤光为谁来”“河汉落酒樽”“水天浮四座”四种不同场景的转换,勾勒出苏轼的传奇人生。渐进式的知识拓展不仅有利于提高受众对传统文化的认知水平,而且能够引领其更好地理解文化的演变和创新,形成更全面深刻的文化认知。具互动性和带来全新感官体验的呈现方式增强了对观众的吸引力,从而推动文化信息更迅速、更广泛、更深度地传播。

4 结语

由中国索引学会数字人文专委会主办的一年一度的中国数字人文年会是一个新兴的学术组织。从2019年敦煌研究院举办文化遗产数字化国际研讨会到2023年末武汉大学信息管理学院承办本次年会,短短五年时间,这一学术组织从无到有,虽尚处于生长的幼年阶段,成长脚步稚拙,却坚韧有力,具有极大的影响力、号召力、向心力和凝聚力,业已成为数字人文学界、业界和相关机构同仁共同的学习型组织、学术交流平台,也是数字人文领域一年工作成绩的展示空间;是彼此激励、相互砥砺的峰会,也是直面问题、探索与争鸣的论坛,是中国数字人文生长的年轮。本届数字人文年会承前启后,自主开放,严而有格,规模超过预期。6场主旨报告较为充分地展现了在AIGC、大语言模型浪潮冲击下中国数字人文研究面临的挑战与应对策略,明确了当下数字人文研究的方向和聚焦的问题。10个分论坛几乎涵盖了当下数字人文产学研各个领域的所有研究方向,论坛活动充满活力,精彩纷呈,既有直面问题的勇气,又有探索未知的锐气;既有成果展示,又有工作反思。29个项目着眼于科研、生产和教学的方方面面,呈现出鲜明的区域性、具象性、创新性、独特性、典型性和应用性。

基于年会所呈现的状况对中国数字人文当前阶段的发展情况进行总结和反思,可归结为以下几点。

(1)从学科发展视角下的研究目的和研究所处阶段来看,目前多数研究仍然处在“验证方法”阶段,即使用数字技术与机器学习方法等进行研究尝试,并与既有研究成果进行比对,以证明数字人文方法的普适性和可信性。

(2)从主要科研产出和技术路径选择来看,在大数据转换为智慧数据的需求和趋势下,智慧数据库、知识库开发与建设问题贯彻数字人文研究的各个领域,反映在本届年会上,即诸多论文及项目均围绕着智慧数据建构、挖掘、共享与使用展开。该类研究涉及大语言模型、知识图谱、本体构建、地理信息系统等多种方法技术,采用RDF、CIDOC CRM、FOAF、DUBLIN CORE等多种国际通用标准。

(3)从实践项目的研究对象与呈现内容来看,现阶段数字人文项目主题向多样化、地方特色化发展。不同地区在进行项目选择时,已经开始更加全面地评估本地区在相关领域的政策支持,考量自身学术背景及学科优势,分析软硬件设施条件,挖掘并充分利用彰显地域特色的文化元素。

(4)从推进范式与方法的角度来看,目前项目多重实践而轻理论,缺乏扎实的、系统性的理论支撑,且多以技术方法为导向而少以研究问题为导向。换言之,目前项目大多关注于现有数字方法和信息技术所能够达到的呈现效果,而在倾听跨学科研究的实际内在需求以及完善研究理论体系方面有所欠缺。

(5)从发展机制来看,受技术、人才、资金、政策、资源等因素影响,数字人文发展的区域差距比较显著,社会机制、学术体系、政府支持、科研人员等多方面协调是缩小差距的重中之重:通过社会机制促进资源合理流动,基于学术体系推进人才培养、交流以及跨地区合作,政府职能部门加大相关政策支持力度等,共同促进不同区域相对均衡又各有侧重地协调发展。

此外,本届年会收到的部分投稿论文或项目存在以数字化技术手段为中心而与数字人文研究完全无关的问题,这从某种程度上说明数字人文作为新兴学科,广大学界和公众对其认知仍然有限,包括数字人文学科概

念与范畴界定、研究内容、研究方法与技术等,数字人文学科的普及仍然有一个漫长的过程。

参考文献

- [1]教育部.教育部关于公布2022年度普通高等学校本科专业备案和审批结果的通知[EB/OL].(2023-04-04)[2024-02-15].https://www.gov.cn/zhengce/zhengceku/2023-04/20/content_5752327.htm.
- [2]武汉大学本科生院.武汉大学数智教育白皮书(数智人才培养篇)[EB/OL].(2023-11-16)[2024-02-15].<https://uc.whu.edu.cn/info/1787/17127.htm>.
- [3]曾蕾,王晓光,范炜.图档博领域的智慧数据及其在数字人文研究中的角色[J].中国图书馆学报,2018,44(1):17-34. DOI:10.13530/j.cnki.jlis.180002.
- [4]ZHAO Q, WANG S. Text as data: a new framework for machine learning and the social sciences. Justin Grimmer, Margaret E. Roberts, and Brandon M. Stewart[J]. Digital scholarship in the humanities, 2022(1):1. DOI:10.1093/lle/fqac068.
- [5]中共中央办公厅,国务院办公厅.关于新时代推进古籍工作的意见[EB/OL].(2022-04-11)[2024-02-15].https://www.gov.cn/zhengce/2022-04/11/content_5684555.htm.
- [6]杨宇婷.数字化赋能文化遗产保护[N].人民日报,2023-08-04(17).
- [7]腾讯研究院.附全文:《中国文化遗产数字化研究报告》重磅发布!“探元计划”2022收官[EB/OL].(2023-03-22)[2024-02-15].<https://mp.weixin.qq.com/s/KpTYUfJ888om5H63ConogQ>.
- [8]孙涛.数字考古:我国历史文化遗产保护的新途径[N].学习时报,2021-10-08(6).

Symbiosis of Digital and Physical: Envisioning the Future of Digital Humanities —An Overview of the 2023 China Digital Humanities Annual Conference

He Qiaoyin Wang Xiaoguang

Abstract The 5th China Digital Humanities Annual Conference(CDH2023), co-hosted by the School of Information Management of Wuhan University and the Digital Humanities Professional Committee of the China Society of Indexers, and organized by the School of Information Management of Wuhan University, the Intelligent Computing Laboratory for Cultural Heritage of Wuhan University, the Big Data Research Institute of Wuhan University, and the Center for Information Resource Research of Wuhan University, was held on 9-10 December 2023 in Guanggu Wuhan. The conference, themed “Symbiosis of Digital and Physical: Envisioning the Future of Digital Humanities”, featured ten thematic sub-forums and garnered extensive attention from both domestic and international higher education institutions and industry sectors. By classifying and reviewing the papers and projects submitted to the conference, as well as analyzing and summarizing the selection of related technology paths and presentation forms, the article aims to explore the current focus and progress of digital humanities research in China, discuss the challenges and coping strategies faced by digital humanities research in China, and forecast future research trends.

Key words digital humanities; symbiosis of digital and physical; conference review