

中国数字人文年会

全景与远景：数字人文与 国家文化数字化战略

——2022 年中国数字人文年会综述

杨泽坤 李少建 凯丽麦

摘 要 由中国人民大学信息资源管理学院、中国人民大学数字人文研究中心与中国索引学会数字人文专业委员会联合主办的 2022 年中国数字人文年会 (CDH 2022) 于 2022 年 11 月 26—27 日在线上召开。会议以“全景与远景：数字人文与国家文化数字化战略”为主题，在国家文化数字化战略背景下，深入研讨中华文化全景呈现与数字人文不断拓展和实现突破的远景蓝图。年会设置了八个专题论坛和两个青年论坛对入选论文进行展示，征集并评选出数字人文优秀论文和项目，并首次开设教育分论坛。文章总结了本届年会的论文、项目、教育分论坛讨论情况，分析中国数字人文研究的现状和热点话题，并展望中国数字人文的未来发展趋势。

关键词 数字人文； 交叉学科； 国家文化数字化战略； 会议综述

分类号 G12； G322

作者简介 杨泽坤，中国人民大学信息资源管理学院讲师，中国人民大学数字人文研究院研究员；李少建，中国人民大学信息资源管理学院档案学专业博士研究生；凯丽麦，中国人民大学信息资源管理学院信息资源管理专业硕士研究生。

0 引言

在信息技术的强力推动下，人类文化正在经历一场剧烈的数字化转型，进入数字新形态。数字人文则是

基金项目：本文得到国家自然科学基金青年项目“数字图书馆情境下基于可解释深度学习的图像检索研究”(项目号:72204255)的资助。

人类文化数字转型过程中必然的学术与文化建设机制。2022年3月,中共中央办公厅、国务院办公厅联合发文,推进实施国家文化数字化战略,为新时代中国数字文化兴起与布局、国家文化大数据建设与应用指示方向。在这一战略实施过程中,数字人文必将推动文化数字化理论与实践研究,助力构建具有中华文化内涵与精神气质的国家数字文化新体系。与此同时,虽然经过20余年的发展,不断经历着他者与自我批判的数字人文,当下仍须以筑牢自身合法性根基为务,从内在学理与外部体制两个层面着力攻坚克难。

在此背景下,由中国人民大学信息资源管理学院、中国人民大学数字人文研究中心(现为数字人文研究院)与中国索引学会数字人文专业委员会联合主办的2022年中国数字人文年会(CDH 2022)于2022年11月26—27日在线上召开。会议以“全景与远景:数字人文与国家文化数字化战略”为主题,深入研讨中华文化全景呈现视域下数字人文的拓展与突破,及以之为方向的远景蓝图。本届年会举办了五场主旨报告、八个专题论坛、两个青年论坛,并首次开设教育分论坛。年会评选出10篇优秀论文和26项优秀数字人文项目。本文将分别总结年会入选论文、优秀项目的特点及教育分论坛的讨论情况,并以之为镜,分析中国数字人文研究的现状、热点和趋势。

1 年会论文总结及特点分析

本届中国数字人文年会共收到摘要投稿135篇,经评审后共向84篇摘要作者发出了论文全文邀请,最终共有48篇正式入选报告环节,其中有10篇文章获评优秀论文。48篇文章按照研究主题共被分至八个分论坛和两个青年学者论坛进行讲演。本文将从内容、方法与技术、作者与所属单位分布三个方面对本届年会入选论文所反映出的中国数字人文领域的部分现状进行呈现。

1.1 研究内容

1.1.1 聚焦古籍处理

本届年会入选论文所涉内容十分广泛,涵盖信息资源管理、文学、历史学、语言学、艺术学、教育学、戏剧与影视学、农学、科技史、党史十余个学科,含纳古籍数智化、近现代档案与文献整理与研究、艺术与文化遗产保护、数字叙事、社会记忆数字保存、知识与语义组织、数字人文教育等内容。其中,涉及古籍处理的论文数量最多,达15篇,处理对象包括《全唐诗》《四库全书》《资治通鉴》、起居注、清实录、二十四史及地方志等。这既是中国数字人文领域迄今积累的体现,也是学界对中共中央办公厅、国务院办公厅印发的《关于推进新时代古籍工作的意见》的积极响应。这些研究提示我们,一方面,因由数字化技术,纸质古籍资源得以在数字空间中呈现与获取,获得传承与保护;知识载体的改变也为进一步“数智化”的知识重组(如从数据库向知识库发展)提供了可能,这种解构与重构的过程使得古籍中的传统文化得以活化。另一方面,综合运用自然语言处理、机器学习与深度学习、语义网、知识图谱等技术,对典籍文本进行深度挖掘,辅助基于古籍的人文研究仍然是当下围绕古籍开展的数字人文研究的主体,并且具体研究对象更为丰富,如防疫知识、野生兽类、水稻品种、孝治文化等。可以预见,数字人文在促进古籍事业繁荣发展、传承和弘扬中华优秀传统文化中的作用将更加突出,可为中华文化的全景呈现、创造性转化、创新性发展做出新的重大贡献。

表 1 入选论文研究大致内容分布

古籍知识库建设		
《史记三家注》引书知识库构建与计算人文研究	刘维菲	南京农业大学
中国古代典籍跨语言词汇底表的构建研究	刘瑞雪 王东波	南京农业大学
基于《二十四史》古现平行语料库的全文检索模型构建研究	梁媛 王东波 黄水清	南京农业大学
数字人文视域下中国易学史知识库的构建与应用	殷展鹏 严承希	中国人民大学
数字人文视域下动物实体数据集构建及分析研究:基于《史记》的探索	林立涛 王东波 刘浏 孟凯	南京农业大学
古籍文本挖掘		
《方志物产》防疫抗疫知识整理与挖掘研究	赵文娟 包平 徐晨飞	南京农业大学 南通大学
《论语注疏》引书主题词的自动挖掘和计量分析	齐月 刘浏 王东波	南京农业大学
大数据视角下的唐诗情感词语与意象关系分析	雷轩铮 黄佩	北京邮电大学
多源数据参照的古典诗歌立体远读法初探——以乾隆四万余首御制诗为例	李斌 何盈学 高策	南京师范大学
明后期至民国时期中国东部大中型野生兽类多样性的变迁:基于地方志与 GIS 技术的量化分析	胡凯津	中山大学
面向数字人文的古代方志水稻品种资源自动整理研究	陆勇 惠富平	南京信息工程大学
意大利中古诗歌及但丁数字化项目	成沫	北京大学
数字人文视角下康熙帝孝治文化分析——以清实录与起居注中间安数据为例	李少建 余敏 罗娜 张佳欣 王文斐 梁继红	中国人民大学
数字人文时代古典文学文献辑佚路径的实验与思考	唐宸	安徽大学
历代科举进士群体的时空演变与网络分析	张强 高劲松 关书朋 夏红玉	华中师范大学 中国人民大学
数字人文视域下古籍的社会网络分析——以《资治通鉴》数字人文知识库为例	常博林 石慧颖 李斌 冯敏萱	南京师范大学
近现代档案、工具书相关研究		
数据库环境下的《同源字典》研究	方树益	中国人民大学
《汉语大字典》研究的聚类和计量分析摘要	黄仁瑄 姜永超	华中科技大学 燕山大学
为民国报刊制作文本基准真相工作中的挑战与思考——以《晶报》为例	谢佳 叶淑敏	德国海德堡大学
可及、表达、共鸣——华侨银信数字人文设计内容开发	黄凯琳 谢欣彤 陈光隼 张思颖	广东财经大学
数字人文视阈下的史学研究趋向与档案馆服务策略初探:以粤海关档案与粤海关外籍税务司研究为中心	詹逸珂 李婧 刘羿伶	中山大学
面向国家文化数字化战略的图书馆创新实践——以张謇学濒危罕见文献数字化保护和利用为例	徐毅 钱智勇 李强	南通大学

续表

近现代档案、工具书相关研究		
机构职官类关系型数据库建置的实践与思考——以“中国共产党组织史专题数据库”为例	李磊	南京大学
数字人文视域下的档案文化产品开发模式研究	周欣然 吴建华	南京大学
电影与艺术		
双重解码:基于数字工具 VIAN 的电影色彩语言分析	朱子彤	北京大学
超越代际:20 世纪九十年代以来中国当代电影的计量电影学考察	李逸 陈涛	南洋理工大学 中国人民大学
新电影史研究中数字人文的介入与反思:以 20 世纪初中日影人社会网络的建构为尝试	江晖 盘大琳	日本东京大学 中山大学
瓣香——中国绘画史的另一形状	王征 陈风明 李静雅	北京交通大学
浅论艺术史研究在数字化转向中的关键问题	齐诗雯 曾达	首都师范大学 中央美术学院
文化遗产保护		
中国编钟数据库建设研究	李梅霞	中国人民大学
非遗传承的社会学习过程分析及数字化技术介入模式研究	张玲燕 王昀	中国美术学院
浅谈数字人文技术赋能文物保护修复	宋平	故宫博物院
数字人文视域下的大运河非遗资源体系本体构建研究	张逸勤 孔嘉 邓三鸿 李佳璇	南京大学
数字叙事		
故事化与游戏化:数字人文的叙事转向、媒介呈现与实现路径	刘珂 刘芮	中国人民大学
事件驱动的数据可视化交互设计:基于“数据叙事”的探索和实践	铁钟 夏翠娟 黄薇 朱永宁 刘梦悦	上海戏剧学院 上海图书馆 中国人民大学 上海交通大学
基于数据叙事的晚清域外游记书写研究	关书朋 张强	上海师范大学 华中师范大学
数字游戏中历史人物形象的传统文化表达	牛雪莹 李艾家	北京师范大学
数字记忆		
网际档案的幽灵:在记忆与遗忘之间	王子云	清华大学
场景再现:面向城市记忆的档案资源开发利用模式与策略	聂勇浩 张俊杰 杜炜欣	中山大学
知识与语义组织		
数字人文视阈下文人画题跋多模态语义分析与语义标注	陈雅	湖州职业技术学院
融合多学科视角的中国古代科技知识数智化——以古代计量知识的分层语义组织框架为例	唐振贵 梁灿兴 罗锦坤	广西财经学院
数字人文视角下多模态知识重构实现路径研究——基于知识供应链模型	杨洋洋 张强	郑州轻工业大学 华中师范大学

续表

数字人文教育		
数字人文视域下的音韵学教学探索	郑林啸	中国人民大学
“新文科”背景下我国高校数字人文教育实践研究	肖可以 秦佳佳 李韞璠	湘潭大学
迈向 2030 元宇宙世代数字人文之戏曲教育数字化战略建议	王台瑞	台湾中国文化大学
数字人文理论		
后疫情时代数字人文的主体建构研究	林恩全 张成良	鲁东大学
数字人文方法论应用及审思——基于 39 例“最佳应对 COVID-19”获奖项目的分析	龙家庆	中国人民大学
全球化社区中中国数字人文的在地化研究	西蒙·马奥尼 付雅明	北京师范大学 上海图书馆

1.1.2 探索红色资源

中共中央总书记习近平在陕西榆林考察时指出：“要充分运用红色资源，深化党史学习教育，赓续红色血脉。”国务院学位委员会、教育部 2022 年印发的《研究生教育学科专业目录(2022 年)》中，“中共党史党建学”以一级学科列入。在此背景下，围绕红色资源开展的数字化建设和数字人文研究正在铺展。本届年会上，南京大学李磊在《机构职官类关系型数据库建置的实践与思考——以“中国共产党组织史专题数据库”为例》一文中尝试对中共组织史专题数据库的核心数据——机构基本信息及领导人名录进行提炼、建立关联模型，以期对长期以来在党史研究尤其是中共组织史研究中相关基础性难题提供支持。这显示出数字人文方法在党史研究领域的应用潜力。

1.1.3 切入社会热点

入选论文中也不乏关注新冠疫情、元宇宙等社会现实与热点的研究。如中国人民大学龙家庆聚焦数字人文方法论的具体应用场景，选取 2020—2022 年间的 39 项国际数字人文“最佳应对 COVID-19”获奖项目作为样本展开分析。南京农业大学、南通大学联合团队的赵文娟、包平、徐晨飞依据物产特征与疾病之间的关联关系构建《方志物产》防疫抗疫知识体系结构；鲁东大学林恩全、张成良探讨了后疫情时代数字人文的主体建构，基于社交媒体中的拟态主体关系链条对数字人文主体的建构进行研究。台湾中国文化大学王台瑞探讨了面向元宇宙世代的戏曲艺术专业教育所需的指标性素养，将元宇宙相关产制技术如 XR(VR/AR/MR) 相关软硬件、容积摄取(Volumetric Capture)、动作摄取(Motion Capture)、立体全息影像(Holographic)等数字工具应用于戏曲教育训练，以推动戏曲素养教育。

1.2 方法与技术

1.2.1 研究方法与技术运用多样

入选论文所用研究方法与技术丰富多样，涵盖地理信息系统、社会网络分析、语义分析、知识图谱、本体建模、机器学习、深度学习、案例分析、文本分析等研究方法，及数据库、LDA 主题模型、Kmeans 聚类算法、TF-

IDF、HanLP、Word2vec、brat 标注平台等技术。其中不乏有单篇文章运用多种方法与技术的情况,如中国人民大学李梅霞所撰的《中国编钟数据库建设研究》提出使用 RDF 资源描述框架、知识图谱、边缘滤波器、Capsule-BIGRU、NLTK、CEDD 等算法和技术对基础性的编钟资源进行深度分析和挖掘,从而形成以编钟价值挖掘为核心的数据库。南京农业大学齐月、刘浏、王东波将 LDA 主题模型、TF-IDF、Word2vec、HanLP、Kmeans 等引入引书研究中。这些文章显示了数字技术对多类型、多模态的人文研究材料进行多元化处理的潜力,材料之中的文化内涵无疑将得到更深度的挖掘。

1.2.2 可视化与交互设计水准较高

入选论文绝大多数都使用了可视化技术对研究结果进行呈现(表2),其中交互设计的水准令人印象深刻。如铁钟、夏翠娟、黄薇、朱永宁、刘梦悦团队《事件驱动的数据可视化交互设计:基于“数据叙事”的探索和实践》一文,以上海图书馆的“上海之源”系列多媒体、数字化展陈项目的数据为基础,根据事件之间的事理逻辑关系,建立非线性的时空关联关系,基于 TouchDesigner 和 Houdini 开发可视化交互设计,并将相关内容呈现于 XR 设备中,构建出能体现出城市历史文化演化关系的数据艺术图谱,实现城市数字记忆的多维数字叙事。

表2 入选论文可视化工具使用情况

可视化内容	可视化工具	出现频次
地理信息	Arc GIS	3
	QGIS	3
	CHGIS	2
社会网络关系	Gephi	5
	UCINET	2
不特定内容	Echarts	5
	Python	4
3D 建模	4Dview	1
	Spatial VR	1
知识图谱	Neo4j3	3
	Protégé	2
视觉电影注释	VIAN	1
交互设计	TouchDesigner	1
	Houdini	1
常规统计	Excel	4

1.3 作者及所属单位分布

1.3.1 分布情况

本次会议入选论文作者共有 111 位,依然显现青年学人积极参与的特点,仅学生作者就达 65 人,占比

58%;其中又以硕士研究生为主(硕士研究生36人、博士研究生25人、本科生4人)。在领军学者的带领下,青年学人已成为学术共同体中的重要力量,这一点也是当下中国数字人文领域的显著特征。

111名作者来自国内外50个机构,包括高校院所、图书馆、博物馆等。作者人数较多的机构有:中国人民大学与南京农业大学均为16人,中山大学8人,南京大学7人,南京师范大学、华中科技大学均为5人,南通大学、上海图书馆、广东财经大学均为4人,北京交通大学、北京师范大学、湘潭大学、上海师范大学均为3人。

本文进一步统计了高校作者所出自的院所,以考察其学科背景。高校作者来自以下学院:信息资源管理学院(或信息管理学院)、文学院、国学院、历史学院、社会学与人类学学院、艺术学院、艺术与 design 学院、美术学院、戏剧学院、音乐学院、外国语学院、语言所、公共管理学院、经济管理学院。归类后各类别学院数量分布情况(图1)一定程度上反映了当前我国数字人文研究群体的学科结构。信管类学院有41所,仍然占据着最大比重,“作为人类社会信息世界的管理者和共建者”^[1],信管学科将与信息密切相关的数字人文科学纳入自身体系有其内在的必然性与适应性。^[2]艺术类学院排在信管类、人文类之后,数量位居第三,引人注目。

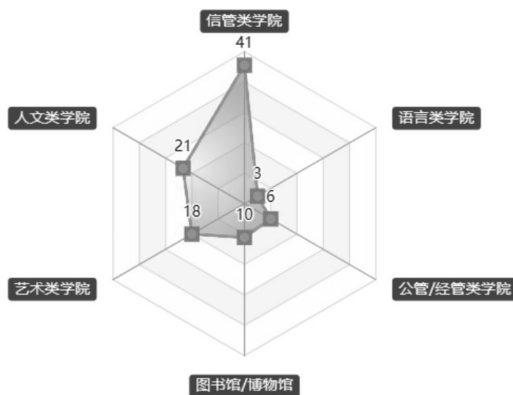


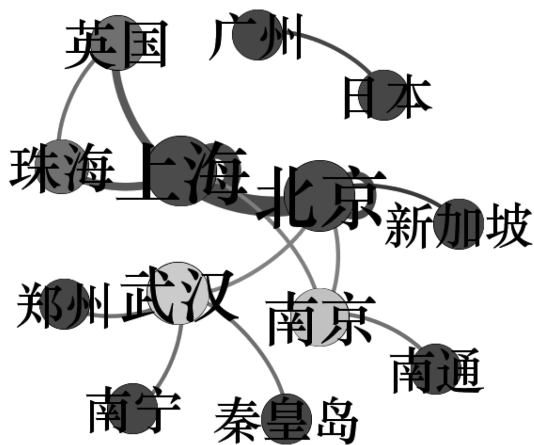
图1 入选论文作者所属院所数量分布

1.3.2 合作研究情况

团队合作是数字人文研究的显著特征之一,而且由于往往需要多资源平台和多领域工具展开资源与方法整合,因此跨校、跨机构、跨地域合作是数字人文研究的常态。入选的48篇论文中合著有36篇,占比75%,跨机构合作有13篇;10篇获奖优秀论文中合著占8篇。从地域来看,参与合作的学者单位分布在国内9个省(直辖市、自治区),其中,北京11家,上海6家,江苏5家(以南京为主),湖北4家(以武汉为主),广东3家(以广州为主)。结合图2来看,北京、上海、南京、武汉仍是国内数字人文研究的重镇,广州则为后起之秀;北京与武汉的学者跨区域合作面最广;上海学者跨区域合作最频繁,其中与北京学者合作次数最多。

与此同时,本次大会入选论文中也有部分是独撰,这些论文可分为几类。第一类是理论性、分析性研究,技术性不突出,所需的技术实践工作较少。如清华大学王子云对网际档案里的数字记忆与人文展开了新的审视与阐释;安徽大学唐宸对数字人文时代古典文学文献辑佚路径展开实验与思考;湖州职业技术学院陈雅基于系统功能语法和视觉语法理论分析,构建了“三层十级”的文人画题跋语义标注框架。第二类是作者使用现成工具,结合自身专业知识即可完成。如北京大学朱子彤使用数字工具VIAN进行了电影色彩语言分析的研究;中山大学胡凯津利用地理信息系统软件ArcGIS重拟了14类兽类动物的分布,结合人口史资料研究当地的

生态变迁。第三类是作者拥有一定的技术能力,可以胜任小规模材料的研究。如北京大学成沫利用 Python 语言自行开发了文本分析软件,针对意大利中古诗歌的格律特色进行优化,目前可进行如重复韵词组、特定韵词查询等操作。



注:原点大小代表学者数量多少;线条代表合作关系,线条的粗细代表合作数量多少。

图2 入选论文地域合作情况

2 年会数字人文项目总结及特征分析

本届年会共征集到 46 个数字人文项目,最终 26 个项目入选优秀项目,其中团队项目 25 个,个人项目 1 个。地域分布上,13 个团队来自北京,8 个团队来自上海,3 个团队来自武汉,1 个团队来自苏州;跨地区合作项目共计 7 个,项目成员来自 16 个城市。机构分布上,共有 35 个机构参与到 25 个团队项目中,包括 20 所高校、6 个商业机构、4 个公共图档博机构、3 个科研机构、2 家出版社、1 个社会机构,跨机构合作项目 6 个。其中,高校仍然是最主要的参与主体,由高校组织的项目共 15 个,其中 4 个在高校之间合作开展。从投稿数量来看,位居前列的有中国人民大学、清华大学、南京大学、武汉大学、西北大学和宁波诺丁汉大学。公共图档博机构中上海图书馆投稿项目最多并全部入选(4 个),在同类型机构中保持了国内领先的地位。商业机构的参与度在本届年会投稿项目中得到了更多的反映,一定程度体现了商业机构的灵活性与数字人文研究开放性、个性化相匹配,在知识服务方面拥有行业优势。社会机构的参与、个人项目的出现更体现了中国数字人文领域参与主体进一步多元化的发展趋势。有一点遗憾在于,本届大会港澳台地区和国际项目代表性有所不足,仅香港科技大学李中清、康文林研究团队提交了“中国历史官员量化数据库(清代)-缙绅录”项目,中西部地区也无项目入选。

从项目类型来看,数字资源建设类(数据库、综合网站)项目共有 18 个,一定程度上反映了基础设施建设依然是中国数字人文领域的发展重心;其他类型包括平台与工具、专题网站、VR 游戏和微信小程序。可视化和视觉传达技术在各项目中发挥着重要作用。

以下将对本届年会的特征进行进一步分析,探求其中反映出的当下中国数字人文领域的部分发展

趋势。

2.1 资源类型与来源

从资源层面看,本届年大会项目既有一贯以来以古籍、文献等为对象的基本历史文化资源的集合与重组,又有以文物、音乐、工艺、网络文化等为对象的创新性资源挖掘,所涉资源类型包括文本、图像、音频、视频、动画和三维模型等。

综合型服务平台仍然多基于文本和图像型资源构建。古联(北京)数字传媒科技有限公司负责建设和运营的“籍合网”是一个古籍整理与数字化综合服务平台,通过整理古籍资源、石刻资源、木版年画、书法作品、甲骨文卜辞和中国字符等内容,建立了十余个数据库,其中“中华经典古籍库”上线资源 5000 余种,累计约 20 亿字。由起承研究院开发的“非物质文化遗产基因”数据库资源类型包含非遗工艺文本和传统纹案图像,囊括 30 类中国传统纹案,搭建了以非遗知识存储和整理为目标的非遗文化多媒体资源库。中国人民大学清末民国社会调查文献研究团队的研究成果“今鉴:清末民国社会调查数据库平台”以清末民国社会调查报告为主要内容,搜集整理了社会调查 4600 余份。由上海外国语大学创建的“中国古籍基础数据应用平台”收集整理了先秦至清末 3.5 万余种传世和出土文献,创建了融合工具平台,实现文本实时分析与可视化、古代汉语语料查询和在线词汇编纂等功能。“吴宝康学术名人知识库”由中国人民大学学生团队建立,通过挖掘名人档案资源,以小专题展现大图景。中国人民大学“北京孔庙”数字记忆项目将繁琐难懂的古代祭孔典籍,通过数字技术转化成具有视听和场景体验感的多媒体数字空间。

音频资源数据库、知识库是本次年会项目的一大亮点。中国数字文化集团推出的“中华音乐数据库平台”提供五大专业音乐典藏数据库服务,包含中华民乐、绝版京剧、中华戏曲、中华曲艺和丝路民乐数据库,是保护和传承民族音乐文化重要的基础性工作。上海图书馆“馆藏唱片知识库”收入 434 种老唱片的 1317 种曲目,并对数据进行了挖掘与关联,为读者提供内容与知识服务。

“‘魂瓶’青瓷数字博物馆”由宁波诺丁汉大学图书馆宁波历史数字人文项目组创建,该项目通过 3D 框架搭建了一座集线上解剖、测量可视化于一体的汉代至两宋时期越窑青瓷数字博物馆平台,包含了文字、图片、视频、动画和 3D 模型等多种类型资源,体现了数字人文项目数据类型多元集成的趋势。

从资源来源的角度来看,本届年会项目还提示了网络原生数字资源在数字人文项目中运用的前景。个人项目“中文梗博物馆”负责人邵泽浩收集了 2000 年以来中文互联网上的一类特殊流行话语——“梗”,创建了 VR 虚拟博物馆。西北大学“秦史蝗”研究组的“中国人口贩运地理信息系统”(CHTGIS)采集了多年来各公益寻亲平台发布的多模块寻亲信息数据。

资源融汇更体现在专门的功能平台上。中山大学信息管理学院的“多维度图像智慧系统”(MISS)基于 LIBRA(Linked Data、IIIF、Big Data、RDF、AI)技术理论,构建了古籍、手稿、文物等藏品图像资源一站式应用生态,可实现不同文化记忆机构图像资源的共建、共享、共融、共生。武汉大学数字人文研究中心与武汉大学文化遗产智能计算实验室联合建设开发的“海外博物馆里的中国”收集了十多家海外博物馆的中国文物数字资源,包括 13 万个文物记录和 20 万张文物图片,种类包括陶瓷器、铜器、玉石器、书法、雕塑和摄影作品,利用关联数据、知识图谱、IIIF 等技术,创建了文物图像交互式大数据资源平台。

2.2 内容主题

本届年会项目主题大体可以分为四类:古籍文博类、城市记忆类、当代文化类和现实关切类。古籍文博类是主要类别,以古籍文献、文物和文化遗产为主要对象。除 2.1 中已提及的项目外,还有如清华大学学生团队开发的“古文智能标引”项目,特点是针对在校各层次学生提供高效的自动标点和专名识别功能。此类别项目的总体特征是工具更为齐全、功能更为完善,可支持多路径的数字人文研究。

有 4 组项目聚焦城市记忆。城市记忆不仅是城市居民实现自身认同、增强城市归属的基石,同时也是重要的学术科研资源。³苏州大学团队建立的“苏州历史文化网”广泛集合了苏州现存的各类型历史文化资源。宁波诺丁汉大学图书馆主导的“宁波城市历史数字人文信息化平台”对宁波城市文化遗产进行了数字化并做全面管理,尤其突出的是古建筑的数字化。中国人民大学、上海大学师生团队合作的“红色夏天 智能航宇”项目聚焦上海近代红色资源,以文旅为重点。南京大学、清华大学、南京艺术学院联合团队创建的“文都时空”文学大数据可视化平台利用本体构建、知识图谱和人工智能等技术深入挖掘南京历朝历代文学文化资源,实现城市记忆中的文学资源智慧化。

本届年会出现了反映现实关切的项目,尤为引人注目。“中国人口贩运地理信息系统”将 1900—2022 年近 8 万条中国人口贩卖数据进行时空可视化表达,为公众进一步认识人口贩卖问题、寻亲信息扩散、相关部门开展治理提供了参考,意义重大。“献礼二十大——历届党代会的信息可视化”项目使得历届党代会的重要信息得以梳理、传播,响应了中共中央关于“用好红色资源”的精神。

聚焦当代文化的项目极具时代特色,可谓当代记忆的重要载体,并尤其体现数字人文项目的个性化趋势。“中文梗博物馆”生动形象地展示了中国网络流行文化之一——“梗”的发展历程和社会背景,为互联网文化研究提供了新的视角和思路。中国人民大学数字人文研究中心学生项目“遇见缪斯·北京”集合了 70 余家北京艺术机构信息,提供艺术活动检索服务,以服务性推动当下的文化潮流。

2.3 参与主体

如前文所述,高校、科研机构之外的主体深度参与数字人文领域建设这一点,充分反映在了本届年会项目中。这些主体又可进一步分为出版机构、商业机构、社会机构、个人。

出版机构方面,中华书局的全资子公司古联(北京)数字传媒有限公司创建的“籍合网”,为传统出版机构建立于自身优势之上的数字转型的成功典范。

商业机构和社会机构主导的项目则体现了数字人文领域在体制外的广阔发展空间。这些项目能将非体制内的研究者与高校、科研机构、图档博等公共机构的研究者灵活地组织在一起,开辟出多元的数字人文项目生长路径。如“中国非物质文化遗产基因数据库”,其发起方起承研究院是联合高校、企业和研究机构开展非遗研究的社会性研究单位。更为典型的是由成都发现时代科技公司开发的“地图书知识库”。该平台源于创始人创建的“观沧海历史地图”公益网站,依赖于志愿者运营团队发展起来,现已成为一个面向数字人文领域的在线协作平台,旨在为科研机构、志愿团队、学者、学生和爱好者提供免费的、即开即用的结构化知识库协作、共享、分析、应用等在线软件云服务。这无疑为体制外、学术界之外的数字人文研究者和爱好者提供了参

与途径,扩充了中国数字人文的力量和根基。

2.4 公众性、公益性

本届年会项目的公众性、公益性取向更为明显。首先,大部分项目都不仅仅面向专业人士,而将向公众展示和传播特定领域知识作为项目目标之一,因此在呈现和功能上交互动性强,充分考虑了受众感受。这方面尤其以本身即为公共机构的上海图书馆的选送项目为典型。“君从何处来”项目对约7万种家谱进行挖掘,并设计了三种展演模式:各朝代的迁徙动态图、迁徙大事件演示和双屏互动与数据筛选。“灵境石语”碑帖艺术展项复原了吴湖帆在《九成宫醴泉铭》中绘制的《九成宫》的场景,观众可以通过沉浸式交互游戏,全景体验唐太宗在九成宫发现醴泉的历史故事。

其次,许多平台都采用了众包和资源征集等方式,扩大了公众参与。例如,武汉大学“珞珈图腾家谱数字化平台”向大众开放,提供在线家谱编修,进行“众包”录入,而得以扩充家谱数据,并普及家族文化。

第三,将公共资源公益开放、服务社会的理念在数字人文建设中得到更多认同,由中国社会科学院、国家图书馆、国家档案局和中国历史研究院牵头,中国社会科学院中国历史研究院近代史研究所承办“抗日战争与近代中日关系文献数据平台”即集中体现了这一点。该项目平台上传各类文献总量已超过5300万页,成为亚洲乃至环太平洋地区资料规模最大、文献种类最全、开放程度最高、受众范围最广的公益性抗战史料平台。其开放、公益的理念使得在边远地区高校师生与大城市的研究者享有同样的文献,实现了“学术平等”,具有重要的示范意义。

3 教育分论坛

国家数字化战略的推行,刺激了相关研究领域和产业对数字人文人才的需求,数字人文教育议题变得更为紧迫,鉴于此,本届年会增设了教育分论坛。中国人民大学信息资源管理学院院长刘越男作为召集人和主持人首先介绍了该学院数字人文教育建设历程和现状。美国肯特州立大学信息学院教授曾蕾作了“全球视野下的数字人文教育”主旨报告,从数字人文的教育模式、课程设置、合作学科、教育管理和行政问题、相关职业、iSchools(信息科学国际联盟)与数字人文六大方面阐述了iSchools数字人文教育的调查结果。来自国内十余所大学和研究机构的16位学者参与了讨论,所涉及议题可归纳为六项,即数字人文的教育类型、思维素养、教育合作、教学体系、人才社会需求以及困境与挑战。

3.1 数字人文教育类型

与会学者认为数字人文教育可以分为三种类型,包括通识教育、专业教育、终身教育。数字人文通识教育旨在面向全体学生进行数字人文基础知识及理念的普及。南京大学艺术学院陈静认为应当持续在本科层次推广数字人文通识教育,并介绍了南京大学本科层次数字人文公选课的情况。清华大学人文学院李飞跃指出,数字人文基本的理论方法和模型应当设置为一门通识课。北京大学图书馆聂华认为,数字人文教育应该主要是通识教育,从支持储备研究型人才的研究生培养迁移到从广度到深度全方位、深层次支持高等教育转

型。中山大学信息管理学院唐琼介绍了中山大学数字人文通识课程的发展情况。上海师范大学人文学院王贺指出,数字人文课程可以设置为人文学科的一门必修课。中国社会科学院赵薇介绍了清华大学日新书院面向本科一、二年级人文学科学生的数字人文实验性课程。

数字人文专业教育,则主要面向特定人文学科学生,针对特定人文学科的研究主题展开。王贺总结了当前(2022年11月)国内数字人文专业学位设置的现状,目前只有中国人民大学、上海师范大学设有专业学位点。陈静认为在研究生阶段,数字人文教育应当回归到专业本身,应当与特定人文专业融合,课程设置应具有专业性。李飞跃认为,虽然数字人文的初级课程应为通识课,但在中高级课程中,应当结合传统人文领域经典问题进行讲授。上海外国语大学图书馆研究员欧阳剑指出,数字人文教育以学科专业研究为主,应当面向学科专业进行人才培养,上海外国语大学的数字人文课程就是主要面向语言学学生进行课程设计。唐琼介绍说,中山大学信息管理学院主要依托文献与文化遗产保护专业硕士学位来推进数字人文教育。武汉大学信息管理学院王晓光指出,当前数字人文教育主要有两条道路,第一是专业化教育,设置数字人文学位;第二是分散在不同人文学科之下,二者是不冲突的,可以共存。

数字人文终身教育则是面向在职者,旨在解决数字人文从业者在实践中遇到的实际问题。聂华指出,2022年,第三届联合国教科文组织世界高等教育大会在西班牙举行,终身教育是会议聚焦的高等教育变革的六个未来方向之一,数字人文教育应当迁移到终身教育和全民教育范畴,支持全社会向学习型社会的转型,而图书馆应当在数字人文终身教育中发挥重要作用。南京师范大学吴静介绍了该校面向南京市100多位处级干部开展数字人文方法和大数据培训课程的基本情况。

中国人民大学冯惠玲提出,数字人文目前尚处于成长阶段,需要保持一段时间的开放性探索,也许不需要统一模式,可既有通识教育、专识方向,也有开放课程等。

3.2 数字人文思维素养

中山大学信息管理学院陈涛指出,数字人文学生的思维素养应当是多样的,不仅需要有人文思维,还应当具有分析思维、计算思维、实验思维等。欧阳剑介绍说,上海外国语大学开设数字人文课程的主要目标是培养学生的数字化研究思维能力。聂华认为数字人文的思维训练不应仅仅是面向数字人文专业学生的,而且应该是渗透进整体人文学科的人才培养之中,数字时代需要学生具备工程思维、计算思维、设计思维。吴静提及南京师范大学数字与人文研究中心更加注重数字方法批判、大数据技术应用、技术治理等对于数字人文应用的各方面透视。云南大学历史与档案学院潘威提到传统人文学科如历史学专业学生面对理工科知识体系和计算思维训练时容易存在抵触情绪。北京大学外国语学院苏祺介绍了北京大学面向人文学科学生开设的计算思维培养课程的基本情况,并提到许多学生在学习编程时比较不适应。王贺也提及人文背景的研究生在做数字人文相关研究的过程中,存在着深入学习兴趣不足以及数据科学训练缺乏等问题,导致实操实作能力有限。赵薇强调数字人文教育追求的是给予学生不断反思、检验、改正的空间,这与人文主义所关心的何为启蒙的价值取向殊途同归。冯惠玲提出,数字人文教育要保持开放、包容、批判、创新、合作的理念,不仅要注重数字技术方法的训练,也要特别注重学生人文与科技素养以及综合学术思维的养成,以育人树人为目标,构建数字人文的教育格局。

3.3 数字人文教育合作

多位与会专家提到数字人文教育实践中合作的重要性。数字人文教育的合作可简要分为高校间的合作及高校与业界的合作两种。关于高校之间的合作,曾蕾指出,iSchool应当积极与其他部门和学校在数字人文领域合作,以满足学生们需求,同时也为iSchool研究人员创造合作机会,并让非iSchool教师和学生接触iSchool的研究和项目。陈静希望中国高校之间能够共享一些数字人文教育资源。陈涛认为中国数字人文联盟可以设置一个虚拟教研室,鼓励数字人文教育工作者在其中共享教育、工具、方法、理论等,这有利于中国数字人文教育体系的贯通。李飞跃指出,不同高校的专业特长不同,应当建立一种跨校机制,在数字人文某些领域利用线上授课模式进行合作教学,共同建立优秀案例集。苏祺认为中文数字人文教育机构可以共同建立慕课平台,让对数字人文感兴趣的学生在平台上进行自主学习。唐琼认为未来可以在数字人文共同体中推动更多的课程建设和教材编写的研讨和共享。吴静认为可以推进广泛的数字人文暑期学校的共同建设。

关于高校和业界的合作,唐琼介绍了中山大学与国家古籍保护中心、广东省档案馆等机构建立的长期合作关系,该项合作主要服务于古籍保护学科建设、地方文献整理、古籍保护学科交流、人才培养以及素养提升等各个领域。

3.4 数字人文教学体系方法

刘越男指出,数字人文理论体系、方法体系、内容体系还远远没有成熟,这为数字人文课程体系的设置带来难度。曾蕾总结目前iSchool的数字人文教学体系,指出当前课程设置通常提供一到两门专门的数字人文课程,第一门课主要介绍数字人文研究中的通用定义、概念和研究方法,第二门高级课程可能会为应用型的、基于项目的数字人文工作提供支持。南京农业大学包平认为需要基于现有理论和方法,尽快形成一套公认的数字人文教材。陈静认为数字人文教育还需要提供丰富多样的教学模式,比如鼓励学生创建小型项目。中国人民大学梁继红指出,数字人文是一种方法论,与传统人文学科已有的方法论存在融通问题,所以应当在本硕博培养方案中把掌握技术工具作为一项重要的教学内容,尤其是在本科层面。聂华认为数字人文教师需要探索新的教育方式,构建新的教学内容,创造新的教学方法,文科人才培养需要强调跨学科属性。潘威介绍了云南大学数字人文教育的情况,表示正在努力将分散在云南大学各个学院的数字人文课程整合起来,形成本硕博数字人文教育体系。南京大学王涛认为数字人文教育的重点是向学生讲授数字人文最核心和最本质的知识,包括技能、理论、方法等。王晓光认为数字人文学科架构设计可以非常多元,需要体现实践性、反思性、多样性。赵薇认为,要重视方法及原理的教学及强调学生的自主学习,要注重渗透检验意识及通过建模来推理、论证的能力,讲授时注重将技术与人文的结合历史化。冯惠玲指出,目前数字人文没有权威的定义和定理,也没有标准答案,这是数字人文发展的必要阶段,因此数字人文教育应当是跨建制的,需要突破传统教育框架去进行教学设计。

3.5 数字人文教育的社会需求

面向社会需求和就业市场是数字人文教育建设的一大重点。曾蕾指出,数字人文课程讲授的技能需要和就业市场联系起来,目前数字人文毕业生最大的雇主是高校图书馆,第二大雇主是大学的数字人文中心。包平认为,在国家振兴的过程中,我们需要数字人文人才来拓展、弘扬传统文化产业,因此数字人文课程体系和培养方案需要重充分了解市场需求。

3.6 数字人文教育的困境与挑战

数字人文教育在学科体系、学科架构设计、学位颁发、课程设置等方面依然存在困境和挑战。刘越男指出,数字人文教育的一大挑战来自外在的学科建制,在我国高校开展数字人文的教育,必须以某一个学科的名义来获得人才培养的授权,数字人文依附于哪一个一级学科更合适,是一个尚未解决的问题。^① 聂华认为,数字人文作为一个新兴的学科,它需要在整个科学体系中确定学术的相对独立性;数字人文应该既是学术分类的名称,又是教学科目。上海大学文化遗产与信息管理学院王丽华指出,从学科建制的角度上来讲,数字人文面临困难,数字人文具体放在哪一个大类的哪一个学科下面,需要形成多方共识,从学科建设的角度来说,硕士教育较易形成突破。王晓光指出,数字人文学位设置存在挑战,是否应当颁发数字人文学位目前还存在相关问题。吴静提出数字人文成为独立学科的意义在于指向人文研究本身,传统的学科建制和学科管理规划使得数字人文教育发展受到限制。冯惠玲认为不应用传统学科框架来界定数字人文。王涛提出,希望高校行政能够突破传统文科管理的方式,在管理层面具备一些新的思维,如跨界合作、革新评价体系等。

4 总结

本届数字人文年会在国家文化数字化战略背景下,较充分地展示了中国数字人文的发展成果,凝聚了数字人文学界的力量,并展现出当下中国数字人文领域的以下特点。

一是聚焦古籍数字化。本届年会上许多论文和项目均围绕着古籍的数字化保护、数字化整理和数字化开发。高校、商业机构和图档博机构将纸质古籍资源通过电子扫描和 OCR 文字识别技术转化为数字资源,利用自然语言处理、机器学习、知识图谱等技术对数字化古籍文本进行知识挖掘与整理,并通过开放数据库、数据可视化、网络平台建设等多种方式向大众展现古籍数字化成果。古籍数字化建设已经实现了全流程打通,成为古籍事业进一步向数智化发展的基础。

二是关注数字人文教育。在数字人文实践已经遍地开花、数字人文人才仍然稀缺的状况下,本届年会特设教育分论坛,以推进在地化语境下中国数字人文教育的探索。参会学者从数字人文教育的类型、素养培育、教育合作、教学体系方法、社会需求、困境与挑战等多个方面探讨了中国数字人文教育的发展思路,凝聚了共识。

^① 2023年4月6日,教育部发布最新的《普通高等学校本科专业目录》,数字人文作为新设专业归入“中国语言文学”类。

三是服务文化数字化战略。《关于推进实施国家文化数字化战略的意见》提出要“建成物理分布、逻辑关联、快速链接、高效搜索、全面共享、重点集成的国家文化大数据体系，文化数字化生产力快速发展，中华文化全景呈现，中华文化数字化成果全民共享、优秀创新成果享誉海内外”。本届年会在中国文化遗产和城市文化记忆两大主题上呈现出许多优秀的论文和项目成果，多位学者探讨了如何在数字人文教育中融入中华文化数字化内容。相信中国数字人文领域将持续助力中华文化数字化，为国家文化大数据体系和中华文化全景呈现的实现贡献力量。

参考文献

- [1] 周文泓,刘静. 数字人文和图书情报与档案管理的双向构建要点研究[J]. 图书与情报,2019(6):101-110.
- [2] KOLTAY T. Library and information science and the digital humanities:perceived and real strengths and weaknesses[J]. Journal of Documentation,2016,72(4):781-792
- [3] 霍艳芳,何思源. 数字人文视阈下城市记忆资源整合与开发路径研究[J]. 档案学研究,2018(1):29-34.

Panorama and Prospect: Digital Humanities and National Strategy for Digitalization of Culture—Overview of the 2022 China Digital Humanities Conference

Yang Zekun Li Shaojian Kai Limai

Abstract Co-hosted by the School of Information Resource Management of Renmin University China, the Research Center for Digital Humanities of Renmin University China, and the Digital Humanities Professional Committee of the China Society of Indexers, the 2022 China Digital Humanities Annual Conference (CDH 2022) was held online from November 26th to 27th, 2022. With the theme of "Panorama and Prospect: Digital Humanities and National Strategy for Digitalization of Culture" and under the background of the National Strategy for Digitalization of Culture, the conference discussed the long-term blueprint for the panoramic presentation of Chinese culture and the continuous expansion and breakthrough of digital humanities in depth. The conference held eight general forums and two youth forums to present the papers, selected and awarded outstanding digital humanities papers and projects, and founded an education sub-forum. This paper summarizes the papers, projects, and the education sub-forum discussions of this annual conference and analyzes the current research trend and major research topics of Chinese digital humanities research.

Key words digital humanities; interdisciplinarity studies; National Strategy for Digitalization of Culture; conference review