

# 国内数字人文的文献计量分析

吴进, 咎栋\*

中国海洋大学图书馆, 山东 青岛

收稿日期: 2023年11月1日; 录用日期: 2023年12月19日; 发布日期: 2023年12月27日

## 摘要

文章利用CiteSpace对中国知网数据库收录的2005~2022年数字人文相关文献进行梳理, 分析国内数字人文的发展脉络、主题演化路径及发展趋势, 进而指出本土资源为国内数字人文研究提供了研究基础和广阔空间, 同时应重视图书情报档案资源的有效利用, 加强领域研究人才的培养, 构建具有自身特色的发展机制, 积极推动数字人文的创新发展。

## 关键词

数字人文, 文献计量, CiteSpace, 研究概况, 发展趋势, 启示

# Bibliometric Analysis of Domestic Digital Humanities

Jin Wu, Dong Zan\*

Library of Ocean University of China, Qingdao Shandong

Received: Nov. 1<sup>st</sup>, 2023; accepted: Dec. 19<sup>th</sup>, 2023; published: Dec. 27<sup>th</sup>, 2023

## Abstract

The article uses CiteSpace to sort out the digital humanities related literature from 2005 to 2022 included in the CNKI, analyze the development of domestic digital humanities, thematic evolution path and development trend, and then point out that the local resources provide a research foundation and broad space for the domestic digital humanities research, and at the same time, we should pay attention to the effective use of library and intelligence archive resources, strengthen the research talents and build a development mechanism with its own characteristics to actively promote the innovation and development of digital humanities.

\*通讯作者。

## Keywords

Digital Humanities, Bibliometric Analysis, CiteSpace, Research Overview, Development Trends, Enlightenment

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

得益于 20 世纪 90 年代以来互联网时代数字经济的蓬勃发展,人文科学领域也迎来新的发展机遇。数字人文将计算机网络技术、数字资源与人文科学研究进行交叉融合,逐步形成文理结合的研究方法。2011 年国际数字人文大会在斯坦福大学举行,同年武汉大学宣布成立国内首家数字人文研究中心,2013 年第 79 届国际图联大会专门设立“数字人文”分会场。国内外学界和业界的广泛关注与参与推动了数字人文研究的蓬勃发展,并使之成为人文社科领域的研究热点之一。2020 年 11 月,教育部发布《新文科宣言》,将 2021 年定为新文科建设元年,全面部署新文科建设,为人文学科的发展擘画了更加广阔的发展空间。作为人文学科的发展趋势之一,数字人文必将迎来高质量发展的新一轮热潮。

21 世纪初期,随着数字图书馆的深入发展,数字人文思想进入中国并逐渐引发了国内学者的关注。目前,该领域研究正处于上升阶段且成果数量日益丰富,已有的文献计量分析研究多以 WOS (Web of Science)数据库收录的文献为研究样本,以国际视角或国内外整体发展趋势研究居多,难以精准把握国内数字人文研究的优势与不足;由于国内数字人文领域研究整体晚于国外,利用国外文献进行研究主题演化及发展趋势分析,易致引导国内研究过度追随国外研究的现象,从而造成研究主题趋同,研究思路缺乏创新的问题。因此,有必要在已有国内外文献计量研究成果的基础上,对国内已发表的相关文献展开深入分析,挖掘具有本土特色的研究主题和实践方法,从而为国内学者的研究选题提供相关参考,亦有助于形成具有自身特色的研究路线和发展思路。

## 2. 研究概况

### 2.1. 数据与方法

文章采用文献计量方法,运用 CiteSpace 可视化工具,对中国知网(CNKI)数据库收录的、发文期间为 2005~2022 年、主题为数字人文的研究文献进行可视化分析,综合关键词共现、发文作者及机构共现、高热度论文、关键词突现检测等维度,探析国内数字人文研究领域的热点变迁及发展趋势,以期为领域内相关研究人员提供有益参考和借鉴。

数据集方面,基于中国知网(CNKI)数据库,以“数字人文”“人文计算”等主题词进行检索,检索出文献 1944 篇。为保证文献数据集的全面性与准确性,手工剔除新闻报道、选题指南等不相关文献,最终得到有效文献 1796 篇,利用 NoteExpress 文献管理软件进行整理、标引,并转换成 CiteSpace 可分析的文件格式,以此作为开展分析的基础数据集。

### 2.2. 发文数量

某一研究主题的学术论文发文情况是该领域学术发展情况的重要体现。通过对数据集年度发文情况的统计发现(见图 1),国内最早关于数字人文主题研究的论文发表于 2005 年,此后至 2015 年的十年间,

关于数字人文的论文发表总量相对较少, 2016 年开始发文量迅速增长, 2017 年首次突破了 50 篇, 并在后续几年中保持持续增长的态势; 与此同时, 观察发文来源期刊可以发现, 来自南大核心、北大核心等权威期刊的年度发文比例均高于普通期刊, 说明领域研究目前处于学界专家引领发展的阶段, 整体进程处于高质量成果积累过程中。近年来, 《再议图书馆发展的十个热门话题》[1]《2021 年度中国图情档学界十大学术热点》[2]业界热点趋势讨论中, 图书馆界高度关注数字人文的发展动向, 这表明数字人文研究依然处于发展的快车道, 相关研究方兴未艾。

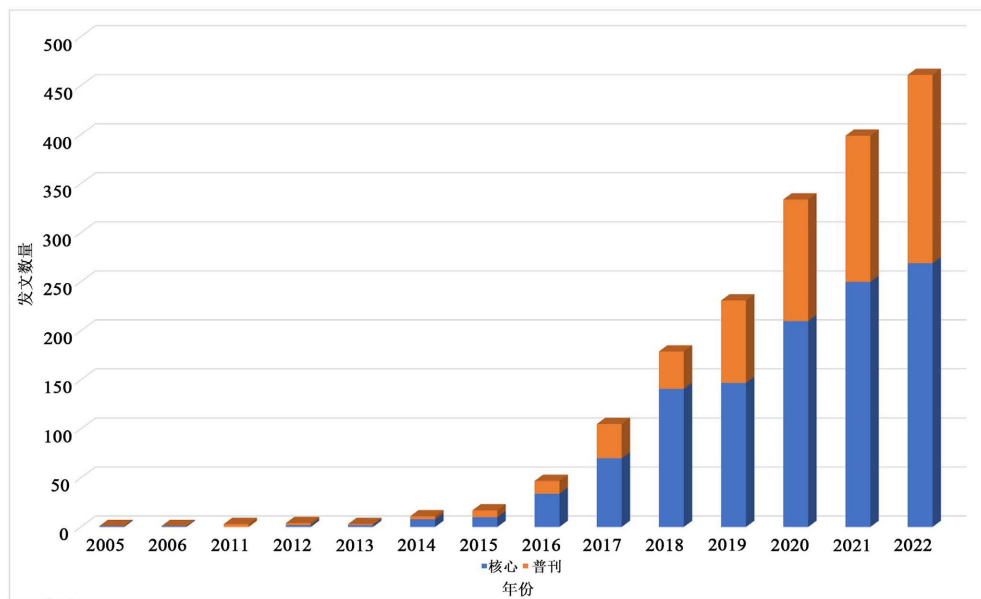


Figure 1. "Digital humanities" publications (2005~2022)

图 1. "数字人文" 发文情况(2005~2022)

### 2.3. 发文学科

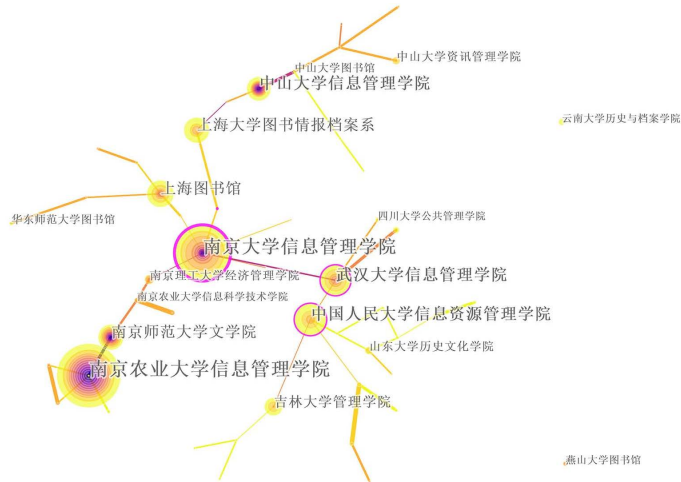
在学科分布方面, 数字人文相关的已发表文献共涉及 30 多个学科, 其中, 图书情报与数字图书馆方向发文占比最高, 比例近 40%; 其次为计算机软件及计算机应用、档案及博物馆, 上述三个学科方向的占比接近 70%; 其余学科涵盖文艺、文学、新闻、历史、考古、哲学等多个研究方向, 并呈现出分布广、比例小的特点。研究方向的跨学科现象多存在于学术方向与图书情报、档案的交叉, 侧面反映出图书情报与档案研究在数字人文学科研究中发挥了重要的支撑和保障作用。

### 2.4. 研究机构

发文机构方面, 发文量排名居于前列的单位主要集中在南京大学、武汉大学、中山大学、中国人民大学等以人文社会科学研究见长的高校, 同时呈现了明显的机构合作特征(研究机构合作情况见图 2), 主要表现在三个方面: 同一机构内不同单位之间的合作, 例如: 中山大学信息管理学院与图书馆、资讯管理学院之间的合作; 同一区域不同机构之间的合作, 例如: 南京大学与南京理工大学、南京师范大学及南京农业大学之间的合作; 另外, 还包括与地域特征无明显关联的机构间合作, 例如: 武汉大学与中国人民大学、山东大学、吉林大学等高等院校之间的合作。

### 2.5. 发文作者

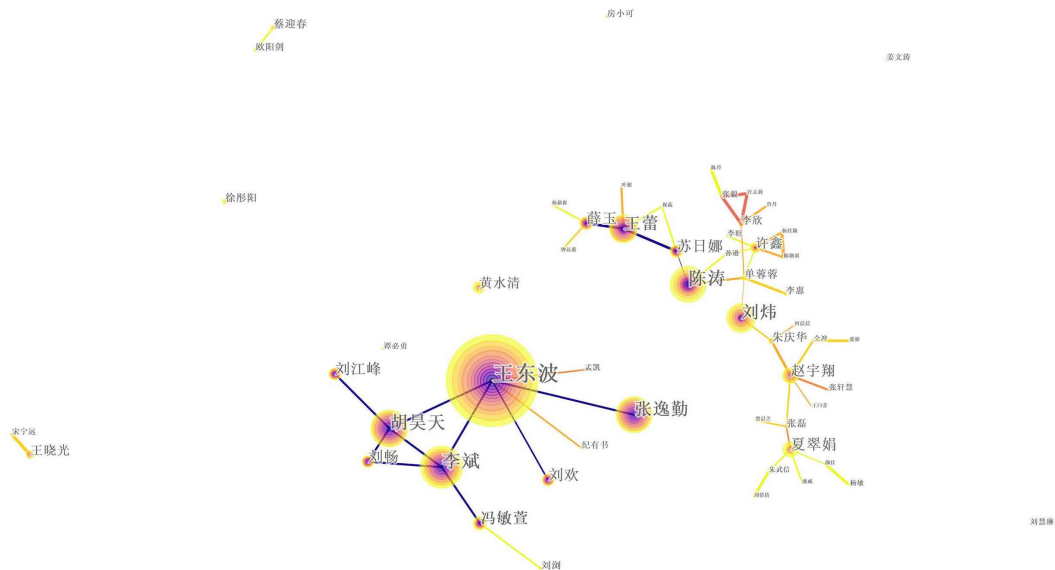
在数字人文领域, 国内期刊发文量超过 10 篇的作者包含王东波(南京农业大学信息科学技术学院)、



**Figure 2.** Collaboration with research institutes  
**图 2.** 研究机构合作情况

夏翠娟(上海图书馆/上海科学技术情报研究所)、刘炜(上海图书馆)、赵宇翔(南京理工大学经济管理学院)、陈涛(上海图书馆)、张卫东(吉林大学管理学院)、黄水清(南京农业大学信息管理学院/南京农业大学人文与社会计算研究中心)、肖鹏(中山大学资讯管理学院)、李斌(南京师范大学文学院)、朱庆华(南京大学信息管理学院)、许鑫(华东师范大学经济与管理学部信息管理系)、邓君(吉林大学管理学院)、鄂丽君(燕山大学图书馆)等。

其中，以王东波、张逸勤、胡昊天、李斌等作者和陈涛、刘炜等作者为代表形成了具有一定规模的作者合作网络(发文作者合作情况见图 3)。在不同合作网络中，相关学者来自不同的学科领域，如数字治理与信息资源管理、计算机软件与应用、语言文字学、新闻传媒、科研管理及图书情报等。与以往单一学科作者的合作研究网络相比，“数字人文”的作者合作网络呈现了明显的学科交叉与多学科融合特征，其跨学科特征主要体现在人文社科领域与图情档案、计算机网络技术之间的融合交叉。



**Figure 3.** Cooperation with authors of publications  
**图 3.** 发文作者合作情况

## 2.6. 研究脉络

关键词聚类分析是一种探索性数据挖掘技术,用于识别和分析特定研究领域中的重要术语主题分类,通过语义算法将数据转换成具有共同属性的集合,从而发现知识领域的研究主题和脉络分布。利用 CiteSpace 对 1796 篇文献的关键词进行聚类,并通过 LLR 算法识别生成聚类,根据聚类规模并结合关键词进行归纳整理,生成关键词聚类知识图谱(见图 4)。按照聚类的规模,依次为新文科建设、档案工作、GIS、图书馆、数字学术、历史档案、世界文学、SIKUBERT、非物质文化遗产、网络文学、开放数据竞赛。国内数字人文领域的研究涉及了方法、模式、内容和参与群体等方面的研究,逐渐形成数字人文领域特有的研究体系和发展脉络,其领域研究的内涵不断丰富和充实,逐步从单一的研究方法向具有包容性、融合性的学科发展,并成为推动人文领域研究的强大动力。

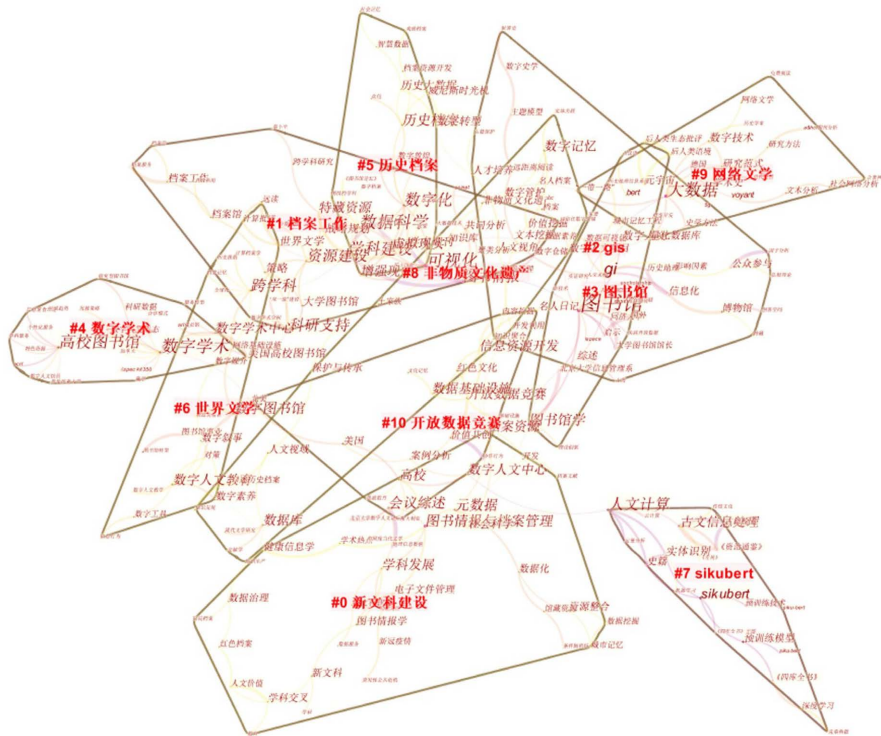


Figure 4. Knowledge graph for keyword clustering based on LLR algorithm  
图 4. 基于 LLR 算法的关键词聚类知识图谱

## 3. 知识演进

关键词是文章核心内容和主要观点的重要表述,对关键词随时间发展的共现情况进行整理及可视化展示有助于了解领域发展研究的主题演进情况。文章以时间分区的视角,可视化关键词共现网络(见图 5),显示了 2012~2022 年期间,出现 5 次以上的关键词,所处的时间段代表该关键词在所属时间段首次出现的年份,排序以关键词共现频率为依据。

综合 2012~2022 年主题演进情况进行分析,2015 年以前国内数字人文领域的研究处于萌芽阶段,主题以宏观研究和理论研究为主;2016 年以后领域研究进入发展阶段,期间《新媒体联盟地平线报告》(2016 博物馆版[3]、2017 图书馆版[4])提及了数字人文技术在博物馆、图书馆等专业领域的结合与应用,强调图书馆界对大数据与数字人文研究的关注以及通过项目实践改变数字人文馆员分析与量化的研究方法,





Figure 5. Co-occurrence time zone map of digital humanities high frequency keywords

图 5. 数字人文领域发文高频关键词共现时区图

拓展馆员在跨学科研究中的技巧。随着一系列具有战略意义的前瞻性报告发布, 业界对数字人文的研究与发展达成共识, 进一步推动领域研究向纵深演化, 逐步分化出人文计算、数字学术、资源数字化与整合、学科建设、人文教育、计算机等多个研究方向。研究参与主体除了人文领域学者外, 图书馆、档案馆、博物馆等文献、档案及实物的收藏管理部门研究人员陆续加入, 其中, 图书馆又分化出高校图书馆和公共图书馆等分支。2020 年以来, 随着《新文科建设宣言》的颁布以及新文科建设战略举措的部署实施, 新文科建设的政策引导与宏观背景为数字人文学科建设带来了新的发展机遇, 数字人文所具有的知识创新本质与学科属性得以彰显[5], 在学科交叉、文化属性、研究范式及发展路径方面都有新的突破与推进。与此同时, 建立数字人文与新文科建设的协同发展机制, 构建数字人文的专业人才培养体系, 相应地成为新的发展方向[6]。

相较于研究主题与研究方法的发展和演进, 过去十年中, 关于数字人文的研究视角发生了诸多变化。数字化技术萌芽初期, 人文学者对数字化技术是否可应用于人文研究持观望态度, 强调对数字化与人文精神的渗透研究, 具体内涵应进行专业性的探索和理性的思考[7]。随着数字技术向各个学科领域的发展与渗透, 人们对数字技术的依赖性不断增强, 随之思维模式越来越技术化和标准化, 学界开始接受并关注数字技术在人文领域的应用[8]。在数字人文发展初期, 学界较多地关注数据、网络、信息资源等保障条件[9], 研究模式方面依赖对国外研究进展和实践模式的借鉴[10] [11]。伴随基础设施建设的规模化, 数字人文开展研究的条件日渐成熟, 数字人文的研究思维和方法开始应用于项目实践中, “数字敦煌” [12]、“民国档案” [13]、“红色记忆” [14]等数字人文项目顺利开展, 社科特色明显的研究机构愈加重视并加入数字人文领域的建设研究队伍, 共同推动学科向更高的目标迈进, 具有中国文化特色的数字人文学科特性, 文理结合的研究范式初步形成。值得一提的是, 在推进基础设施建设的过程中, 图书馆、博物馆及档案馆凭借在资源建设、管理、服务及研究等方面的专业表现, 巩固夯实了其在数字人文领域的重要地位[15]。

#### 4. 发展趋势分析

发展趋势通常指近期突现的研究主题或者在过去已经出现,但是在近期关注度突然提升的研究热点。对这些主题进行关注和跟踪,有助于了解领域研究的发展动向,从而更好地把握研究方向。文章将从两个方面对发展趋势进行分析:一是根据 2012~2022 的关键词突现选择相应的文献,二是综合文献的被引及下载情况提取高热度文献,通过对这两类文献进行分析,得出研究主题的发展走势。

利用 CiteSpace 中的关键词突发功能得出 2012~2022 年期间的关键词突发情况(见图 6),近三年突现值明显的关键词主要是“数字人文服务”“档案学”“档案资源”。早期的数字人文研究中,学术界多聚焦于数字人文技术的发展,而忽略了作为研究主体——人文学者的重要角色[16]。对人文学者参与数字人文信息需求和参与行为进行调研分析,有助于全面理解人文学者复杂的信息行为,为理解数字人文的发展趋势提供新的研究思路;另一方面,对数字人文服务的研究主要来自于图书馆界,结合数字人文自身发展的需要以及图书馆空间再造与服务转型的内在需求,图书馆界在开展数字人文咨询、数据管理、社交协作、信息素养培训[17]等数字人文服务方面开展了大量的探索和实践。在全方位提升空间智慧服务水平的同时,也不断加深与数字人文发展之间的联系,进一步巩固和加强图书馆参与数字人文的地位。2022 年,“档案”主题处于中心热点地位,这表明数字人文研究继续拓展、深化,展现了数字人文研究的旺盛生命力,也提示了档案学与数字人文融合研究还存在继续挖掘研究的空间。



Figure 6. Digital humanities keywords image

图 6. 数字人文关键词图像情况

文献出版后近年的引用次数可表明该主题的研究热度,而短时间的下载量亦可表示最近出版文献的受关注度[18]。基于被引频次和文献下载量,文章收集整理 2019~2022 年期间发表的年度关注度最高的

15 篇文献, 抽取关键词如表 1 所示。对比四年的热点主题, 人文学者、新文科、图书情报与档案管理以及数字化建设、知识组织相关的信息服务在业内一直保持着较高的关注度, 预示着数字人文未来的研究将更加关注研究主体并围绕需求继续加强服务的建设和完善。此外, 红色档案及文化资源建设在近年受关注度颇高, 这与建党百年及 2022 年以来开展党史主题教育有着密切联系。新文科建设广受关注, 与国家的“四新”战略及 2021 年发布的新文科宣言密切相关, 未来数字人文与新文科战略的结合将推动进一步发展, 并成为业界所持续关注热点。古籍数字化与整理一直保持高热度, 这与古籍所承载的文化息息相关, 通过整理古文献, 可以挖掘中华民族传统文化的精华, 让人们在现代化进程中把握民族发展脉络, 并参与世界文化交流。

**Table 1.** High heat keywords in digital humanities 2019~2022

**表 1.** 2019~2022 年数字人文领域高热度文献关键词

2019	2020	2021	2022
高校图书馆	图书馆	图书馆	知识图谱
人文学者	档案学	数字档案馆	文化遗产
知识图谱	古籍数字化	新文科	古籍整理与研究
学术名人	图书情报与档案管理	档案数据	知识发现
档案资源	红色档案	红色文化信息资源	可视化
知识组织	知识服务	古籍数字化	高校图书馆
数据组织	知识图谱	服务模式	数字学术

## 5. 结论与启示

### 5.1. 本土资源奠定发展基础, 提供发展空间

回顾国内数字人文的发展历程可以发现, 虽然国内研究起步晚于国外, 但是我国特色鲜明的历史文化为数字人文领域研究提供了丰富的物质文化遗产和记忆源泉, 为国内数字人文研究的崛起奠定了坚实的资源基础。具有地方特色的历史地理信息和与红色记忆相融合的文化研究, 近年来更展示出强劲的生命力和发展动力, 应充分重视这些具有民族鲜明特征的文化主题的挖掘与开发; 积极发展既具备人文情怀, 又富含实践创新价值的数字人文项目; 重视原理、价值、思维、模式等方法提炼总结与理论体系构建; 在呈现方式方面, 尝试融入大众易接受、互动体验性强的设计理念。在挖掘数字人文项目学术研究价值的同时, 扩大社会参与度和影响力, 为数字人文的可持续发展注入后续动力。

### 5.2. 重视图书情报档案资源的有效利用

图书馆多年来通过开展信息服务, 在用户沟通、平台构建、知识组织及数据发现等方面积累了丰富的经验, 为数字人文研究平台的搭建提供了扎实的支撑条件。高校图书馆坚持开设信息素养课程, 帮助读者提升信息素养和文献利用水平, 有效弥补了人文学者在资源、工具运用等方面的短板。近年来, 数字图书馆的快速发展为数字人文的研究和学科建设提供了丰富的素材和广阔的发展空间, 在信息量飞速增长而数据需求急剧提升的过程中, 图书馆发挥的作用不仅限于资源服务、数据库建设与信息服务。随着数字图书馆向智慧图书馆的过度与发展, 图书馆在资源建设、信息组织及人员队伍等方面均具备了更加明显的优势。图书馆界亦应以研究者的姿态积极参与和助推数字人文的研究进程, 积极发挥知识组织者的作用, 实现产学研的有效联动, 助推智慧图书馆与数字人文的健康有序发展; 应进一步优化完善馆



藏资源的规划、组织及呈现, 并应重视自身社会教育与知识传播属性的发挥, 更好的发挥整合资源、挖掘内容, 融通知识的作用, 实现自身的转型发展价值彰显。

### 5.3. 加强领域研究人才的培养和储备

数字人文研究的跨学科特征需要复合型人才的不输入以形成发展的持续动力。人才培养方面应既注重人文学科背景的专业性, 又强化数字方法及工具的掌握与运用。在具备扎实的专业知识的基础上不断提升工具利用和数据分析的实操能力, 从而实现新技术与人文学术研究之间的融合与发展。目前, 我国在数字人文相关的专业设置及人才培养方面仍缺乏成熟的培养体系。相比而言, 国外在学科建设、课程设计、培养途径等方面已经形成相对完善的发展机制, 可参考和引进国外数字人文在人才培养模式和实践方面的经验, 进一步健全和完善我国的人才培养机制。

### 5.4. 构建具有自身特色的发展机制

目前, 国内积极开展开放数据共享项目建设和研究, 在扩大数字人文研究影响力方面已经开展了一些尝试和探索, 例如开放数据竞赛[19], 高校数据驱动创新研究大赛[20]等, 鼓励不同行业、不同领域的数据学者加强大数据、人工智能、虚拟现实、云计算、数据建模等研究方法与实践手段在社会科学领域的推广普及与研究应用。但我们仍需借鉴国外数字人文中心项目制运行方式在档案管理和历史记忆研究方面取得良好效果的成熟经验, 并与我国文化特色与国情相结合, 建立并形成具有自身特色的发展机制, 从而在扩大数字人文影响力、提升项目研究成效方面发挥更加积极的作用。

## 参考文献

- [1] 吴建中. 再议图书馆发展的十个热门话题[J]. 中国图书馆学报, 2017, 43(4): 4-17.
- [2] 2021 年度中国图情档学界十大学术热点[J]. 情报资料工作, 2022, 43(1): 5-12.
- [3] A·弗里曼, S·亚当斯·贝克尔, M·卡明斯, 等. 新媒体联盟地平线报告: 2016 博物馆版[J]. 开放学习研究, 2016(5): 1-13.
- [4] Adams, B.S., Cummins, M., Davis, A., 等. 新媒体联盟地平线报告: 2017 图书馆版[J]. 图书情报工作, 2018, 62(3): 114-152.
- [5] 王丽华, 刘炜. 助力与借力: 数字人文与新文科建设[J]. 南京社会科学, 2021(7): 130-138.
- [6] 周毅, 李卓卓. 新文科建设的理路与设计[J]. 中国大学教学, 2019(6): 52-59.
- [7] 廖祥忠. “超越逻辑”: 数字人文的时代特征[J]. 现代传播, 2005(6): 23-25.
- [8] 李巧明, 王晓光. 跨学科视角下数字人文研究中心的组织与运作[J]. 数字图书馆论坛, 2013(3): 26-31.
- [9] 夏翠娟. 面向人文研究的“数据基础设施”建设——试论图书馆学对数字人文的方法论贡献[J]. 中国图书馆学报, 2020, 46(3): 24-37.
- [10] 宋玲玲. 大数据时代图书馆数字学术中心服务探析——以北美高校图书馆为例[J]. 图书馆理论与实践, 2018(8): 92-97.
- [11] 缪军. 美国高校图书馆的典型数字学术服务与启示[J]. 图书与情报, 2017(4): 34-37.
- [12] 李思琪. 历史档案资源的数字化开发: 数字敦煌项目的启示[J]. 山东档案, 2020(4): 6-9.
- [13] 陈宇. 民国档案数字人文服务平台建设思路研究[J]. 北京档案, 2021(8): 28-31.
- [14] 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.
- [15] 周文杰. 图情档学科发展的数据化趋向解析——基于 2019 年度学术热点的系统性文献调查[J]. 情报资料工作, 2020, 41(2): 20-30.
- [16] 董舞艺, 梁兴堃. 中国人文学者参与数字人文动机的二元结构及行为路径[J]. 中国图书馆学报, 2019, 45(4): 86-103.
- [17] 唐燕, 刘小榕, 李健. 智慧图书馆空间再造与数字人文服务创新研究[J]. 图书馆, 2020(5): 74-80.

- 
- [18] 司红运, 施建刚, 陈进道, 等. 从《中国人口·资源与环境》审视国内的可持续发展研究——主题脉络、知识演进与新兴热点[J]. 中国人口·资源与环境, 2019, 29(7): 166-176.
- [19] 赵宇翔, 张妍, 仝冲, 等. 面向价值共创的数字人文类开放数据竞赛行动者协作行为研究[J]. 情报理论与实践, 2022, 45(1): 112-119.
- [20] 崔海媛, 罗鹏程, 赵静茹, 等. 数据驱动研究范式和一流高校数据服务支撑体系研究——首届全国高校数据驱动创新研究大赛综述[J]. 大学图书馆学报, 2018, 36(6): 28-36.