

数字人文评价:学科性、专业性、技术性

蔡迎春

摘要 数字人文作为 21 世纪的一个现象级国际话语,其发展深度、广度与速度均在显著提升,同时数字人文评价也开始进入研究者的视野,成为数字人文领域的新课题。本文结合数字人文研究的特点和发展阶段,从学科性、专业性和技术性三个维度构建与探讨数字人文评价问题:学科性是核心要素,评价数字人文成果最终要看它是否推动了人文领域的知识创新和学术创新;专业性是关键要素,图情档学科为数字人文提供数据基础设施,并利用信息组织、信息检索、数据分析等专业知识为数字人文提供方法论指导,也是连接学科和技术的媒介和纽带;技术性是必备要素,信息技术、自然语言处理技术、语义网技术、文本挖掘技术等应用于人文数据分析中,助力人文领域的学术创新和知识创新。图情档领域应发挥专业知识优势、技术应用优势、专业人才优势和专业组织优势,加强跨学科交流与合作,推动人文学科研究的理论创新、知识创新、学科创新和方法创新;同时,借鉴已有评价研究成果和评价实践经验,引领开展数字人文成果的评价。表 1。参考文献 19。

关键词 数字人文评价 图情档学科 跨学科研究 学科性 专业性 技术性

分类号 G250

Digital Humanities Evaluation: Discipline, Specialty and Technology

CAI Yingchun

ABSTRACT

As a phenomenal international discourse in the 21st century, the depth, breadth and speed of digital humanities have been significantly improved. At the same time, the evaluation of digital humanities has begun to enter the vision of researchers and become a new topic in the digital humanities.

It can be found from the related research of digital humanities at home and abroad that the research mainly embodies the characteristics of discipline, specialty and technology. From the overall development trend, digital humanities projects can be divided into three stages: primary stage, intermediate stage and advanced stage. Combined with the characteristics and development stages of digital humanities research, this paper constructs and discusses the evaluation of digital humanities from three dimensions of discipline, professionalism and technicality. Discipline is the core element, and the evaluation of digital humanities achievements depends on whether it promotes the practical and academic innovation in the field of humanities. Professionalism is the key element. Library, information and archives science builds basic data facilities for digital humanities and provides methodological guidance for digital humanities by using its professional knowledge of information organization, retrieval and analysis. It is also the medium and link connecting disciplines and technologies. Technology is an essential element. Information technology, natural language processing technology, semantic web technology and text mining technology are applied to

通信作者:蔡迎春,Email:cyingc@shisu.edu.cn,ORCID:0000-0002-7736-2267(Correspondence should be addressed to CAI Yingchun, Email: cyingc@shisu.edu.cn, ORCID: 0000-0002-7736-2267)

human data analysis to help academic innovation and knowledge innovation in the field of humanities.

In the field of library, information and archives, we should give full play to the advantages of professional knowledge, technology application, professional talents and professional organizations, strengthen interdisciplinary exchanges and cooperation, and help promote theoretical innovation, knowledge innovation, discipline innovation and method innovation in the field of humanities. At the same time, we should learn from the existing evaluation research results and practical experience to lead the evaluation of digital human achievements. 1 tab. 19 refs.

KEY WORDS

Digital humanities evaluation. Library, information and archives science. Interdisciplinary research. Discipline. Specialty. Technology.

0 引言

数字人文作为21世纪的一个现象级国际话语,其深度、广度、速度与温度在数字技术、工具和方法快速发展的背景下均得以显著提升。数字人文适应了整个社会数字化转型和变革的趋势,无论是理论建构层面还是技术应用层面,无论是研究领域视角、研究机构视角还是研究成果视角,都已经拥有较为完备的跨学科研究体系,在实践中出现了一批以高校为依托的数字人文研究机构和以图书馆为依托的数据基础设施建设机构,并产生了较有影响的实践成果和学术成果。

数字人文源于“人文计算”,以人文数据为基础,以计算技术作为有效的辅助和服务手段,帮助研究者提升对人文语料的处理效率。在与传统人文学科思想的交汇中,“人文计算”从单纯的工具方法逐渐向信息技术与人文学科的跨学科融合过渡,并进一步发展成为凝聚多元领域且具有独立的专业实践、研究方法、研究标准和理论价值的新的学科——数字人文^[1]。数字人文的产生与发展,既是计算机技术和方法向新研究范式演进的过程,也是传统人文学科触角向外延伸的现实需要。在我国,数字人文作为一门新兴学科,更是助推新文科发展不可或缺的新生力量。

随着数字人文项目和研究成果的增多,对

数字人文成果进行评价开始进入研究者视野,如何评定数字人文成果成为学术领域的新课题。数字人文的产生和发展依靠三个维度:学科性、专业性、技术性。从学科性来看,数字人文具有人文学科特质和跨学科特征;从专业性来看,图情档学科的理论和方法是支撑数字人文研究的专业性基础;从技术性角度看,数字技术与智能分析工具的开发与应用是推动数字人文发展的必备因素。以上这些特征也使得学科性、专业性和技术性成为数字人文评价的三个主要维度。

1 数字人文的三个维度和阶段划分

1.1 数字人文的三个维度:学科性、专业性、技术性

为了更清晰地展现数字人文的发展脉络和发展特点,本文以“数字人文”“人文计算”和“Digital Humanity”“Digital Humanities”“Human Computing”作为主题词,分别在《中国学术期刊(网络版)》数据库(CAJD)和Web of Science(WoS)核心合集数据库进行检索(截至时间:2020-12-31),经筛选最终获得国内外相关文献1180篇和10969篇。通过深入分析国内外数字人文相关研究成果的主题,总结发现国内外数字人文研究的特点。国外方面的研究主题主要集中在三个方面:一是资源的数字化建设与维

护,主要是将数字化文本、数字资源保存、数字仿真等相关技术应用于计算机地图、虚拟图像而形成相关成果;二是图情档领域参与数字人文的探索性研究,主要包括信息可视化、信息检索、信息服务等内容;三是数字人文相关技术研究,包括数字化技术、可视化技术、信息技术等。国内研究主题也主要集中在三个方面:一是以人文项目为案例的数字化转型探索与研究,例如古籍、民国文献、徽州文书等特藏资源整理与分析项目;二是数字人文概念的界定以及图情档领域在数字人文背景下的发展趋势及服务转型;三是数字人文技术、工具、平台的研究、开发和利用等。

从以上对国内外数字人文相关研究的总结可以发现,国外和国内的数字人文研究有相通之处,即主要体现三个方面的特点:学科性、专业性、技术性。

(1)学科性:数字人文的核心是“人文”,决定了其研究主题的人文学科性质,这也是其核心特质。从中外数字人文的项目来看,每个项目都有相对集中的人文社科主题,或语言、或文学、或历史、或哲学、或艺术,或兼而有之,这是中外数字人文项目的共性,也是数字人文立足的根本。从目前已有的数字人文项目和研究成果来看,虽然早期是图情档领域在主导数字人文的项目实践,发表相关的研究成果,但项目的核心内容和主题都是人文内核。

跨学科是数字人文的最本质属性。国外数字人文的学科交叉深度要强于国内,国外数字人文相关研究主要汇集了人文科学、社会科学、计算机科学、地理信息学等多领域的学者,跨学科性非常明显,研究范围广泛,研究成果互相融合。国内数字人文已有十余年的发展历史,早期的数字人文项目主要由图情档领域倡导引领,相关成果也主要刊登在图情档期刊上。近年来,数字人文开始受到人文社科学者的关注,综合性和理论性的人文社科期刊和报纸如《中国社会科学》《文艺理论与批评》《理论月刊》《中国社会科学报》也关注并刊登了相关研究成果,这说明人文社科领域开始逐渐接受数字人

文,也说明数字推动下的人文学科在研究方法和思想理论方面正悄然发生变革。

目前,我国正处于新文科建设的语境中,“数字人文生态”的提出促使我们从整体性视角来审视当前的数字人文实践。对于新文科背景下的人文学者而言,一方面需要在了解数字人文的基础上,有突破传统人文研究的勇气和能力,以积极的态度寻求通过数字人文平台或相关项目的带动,促进人文研究视角向更为广阔的空间和场景延伸;另一方面也需要具备跨学科合作精神,善于从传统人文研究中去思考数字人文的发展路径,以及数字技术与人文研究进一步融合发展的可能^[2],实现向数字人文学者的转型。数字人文不仅是在某个领域的简单线性应用,或跨领域的不同层面的应用,而且是形成多维度、立体化的网状结构,推动形成各学科领域的数字人文学术共同体^[3]。

(2)专业性:数字人文的实质是利用数字的理念和数据分析技术对文本中的相关词及词间关系进行深度揭示,实现超时空、多维度的共现,使人文学者能够突破原有平面的、文本的视角看待问题,发现新的研究视角并启迪思维。在数字人文项目的策划、组织、实施过程中,图书馆学对于文献的揭示、标引和对知识组织的认知,起到了非常关键的作用。文献标引技术所揭示的知识规则,可以说是构建人类知识体系的基础,也是实现知识发现新方案的基础。揭示标引词之间的内在关联,最终实现对海量文献的分析与处理,已大大突破了传统的个体有限生命和记忆所承受的能力,而自动标引的实现及其在图情档领域的实践与应用,大大提高了文献标引的效率和质量。同时,基于非受控词表,以自然语言为基础,共词分析法、语义网技术和本体方法对人文领域知识体系的有效揭示,也为数字人文研究构筑了广阔的发展空间和前景^[4]。

图书馆丰富的文献信息资源和专业的基础数据服务平台,是数字人文研究中精准数据分析和得出可信研究结论的前提,也是跨学科研

究的保障。因此,图书馆参与数字人文研究的关键是要找准定位,不论是理论层面还是实践层面,都需要重新审视“文献”这个概念,尤其是文献在新的媒介环境下表现出的新的数字化和数据化形态,进而发展成为智慧数据^[5]的过程。发掘传统文献在新的数字媒体环境下的全新变化和潜力,以及利用标引与检索等图书馆学专业理论知识、方法和经验为数字人文提供资源和平台,是图书馆学专业对数字人文发展的两个有力支撑。

(3) 技术性:数字人文利用数字的方法和技术带来了人文研究过程中知识生产和分析的新认识论和新方法论,即知识生产新形态^[6]。数字人文主要以工具的创新和方法的变革为肇始,技术、方法和平台的研发可以说是其研究和实践的物质基础。开发针对人文学者的多样性数字人文工具,会吸引更多人文学者参与到数字人文研究中来。

目前,数字人文涉及的相关技术和方法极其丰富,主要包括计算机技术、信息技术、多媒体技术、自然语言处理技术、语义网技术、文本挖掘技术、GIS技术,以及可视化、VR/AR等数字技术。数字技术将人文学者、资源与媒介紧密联系起来,在人文领域“大显身手”,不断丰富人文研究的视角和方法,努力寻找数字技术与人文学科的平衡点。因而,对于传统人文研究而言,数字人文“可以有效解决学术之间独立和无关联问题”,借助数字技术具有的多样性、包容性和延展性,通过计算、分析、可视化等手段改变或重塑人文知识,为人文学者提供更多宏观性、规律性、差异性和趋势性的研究方法,促进人文学科领域实现“轮廓重绘”。而对于数字技术来说,在注入人文价值观和分析方法之后,其功能则更加强大、多元,可以处理跨媒介、语言、地点、历史的不同问题,引导、扩展甚至改变人文学者的研究习惯,成为助推数字人文发展的重要手段。

数据管理、信息可视化、人机交互技术、机器学习技术等数字人文相关技术和研究方法在

人文、社会科学以及自然科学、计算机领域的不断拓展和创新,为数字人文研究奠定了基础,提供了研究所需的素材、工具和交流平台,逐渐形成“问题跨界、学科融合、工具共通”的特征。

新文科驱动下的人文学者,开始以数字人文为“器”、以数字人文为“思”,主动学习数字人文新技术,把数字人文研究的新范式嫁接在传统研究方法上,并将其作为增益人文研究的引擎^[7]。以大数据、数据管理或数字技术为基础的数字人文学科的建设,以及在人文学科的各分支领域加强人文学者的数字人文相关技术、工具与方法的应用培训,可以让更多的人文学者进入这个领域,也可以让进入这个领域的人文学者产出传统研究方法无法获得的成果。

1.2 数字人文发展的阶段性

数字人文的产生,源自研究者探索和解决如何借助计算机自动输入文本这一问题。1949年,意大利耶稣会学者罗贝托·博萨(Roberto Busa)和IBM合作,使用自动化技术编制了《托马斯著作索引》,首次创造了由计算机生成的托马斯·阿奎那作品索引^[8]。该索引开创了人文领域运用计算机之先河,同时也为图书馆领域参与数字人文研究奠定了专业基础。

可以看出,学者偶然将计算机技术与图书馆学的文献整理理念相结合,形成了最初的数字人文形态和成果。随着社会的发展,知识不断更新,各行各业产生的文献和信息数量日益增多,信息技术被广泛地应用到人文领域文献资料的整理和研究中,并随着文献整理理念的变化不断更新和升级。

在人文领域,对于文献的整理从原始的人工分类、索引、排架、汇编等,发展到机读编目、纸质文献数字化及建成专题资源库,再进一步发展到数据化的文本资源库。对于人文文献信息的利用,从最初的人工检索发展到计算机自动检索,再进一步发展到数字化和全文检索,再发展到目前的文本分析和挖掘、智能分析和计算等。而人文领域本身,随着数字化转型,也在

不断开拓新的研究领域,如文化遗产的数字化保存问题、数据资产问题等。从目前的发展过程来看,数字人文项目可分为三个发展阶段。

(1)初级阶段:构建数字人文基础设施,分不同人文专题完成基础数据和平台的建设。这是数字人文项目建设和研究的基础。上文提到的《托马斯著作索引》,计算机自动编制的前提是具备全面、准确的基础数据,否则也无法编制正确的索引。该阶段也分不同时期,早期主要是开发平台,完成纸质文献的数字化和数据化的组织整理;随着整个社会数字化转型的深入,也要搜集和整理原生的数字化人文资源,不断增加人文数据数量,充实基础设施建设。该阶段是动态的过程,要随着不同领域文献和信息的发展不断更新和补充基础数据。

(2)中级阶段:利用基础数据和信息技术进行数据的自动化处理,或进行深度的数据分析和智能计算,数据可视化和知识图谱呈现是这个阶段采用的主要方法。该阶段的分析成果或与传统的人文研究互证,或辅助传统的人文研究。该阶段以前期建设的人文数据为基础,也为更高级的人文研究奠定基础,是迈向高级阶段的必由之路。这是目前大多数数字人文项目和研究所处的阶段。

(3)高级阶段:人文学者深度参与数据分析或智能计算的过程,利用分析和计算的结果,产生创新性的成果。数字人文真正融入人文社科领域的研究,能帮助人文学者取得传统方法无法取得的突破,推动人文社科领域的学术创新、思想创新和知识创新,弥补传统人文研究的不足。这应该是数字人文发展的高级目标。而且,随着整个社会数字化转型的深入,人文研究范式的数字化转型将会成为常态,“数字学术”将成为未来的发展方向。

以上三个阶段都由人文学者、图情档专业人士和技术研发应用人员共同合作完成。图情档的专业文献搜集、整理和数字化,以及通过技术人员建成人文数据整理平台及其他基础设施,是数字人文项目的基础,也是整个项目继续

推进的前提。在国内,目前主要以图情档人为主,并联合人文学者,利用人文基础数据,进行数据分析和计算,辅助人文领域的研究,并逐渐引起更多人文学者的关注。人文学者开始逐渐认识到“数字”对于人文研究的作用和影响,开始深度参与其中,也逐渐产出获得人文领域认可的学术成果,这是数字人文项目评定的关键。对数字人文项目或成果进行阶段性划分和认识,是进行科学评价的前提和基础。

2 数字人文评价的三个维度:学科性、专业性、技术性

数字人文自产生之日起,就受到业界和学界的广泛关注,也有来自不同领域的不同声音。积极参与者有之——图情档领域积极推动并引领;质疑者有之——人文领域最初对此认可度不高;观望者有之——更多的人认为它仅仅是工具,是比文献数字化更高一级的工具。随着数字人文项目和成果的增多,语言、文学、历史、艺术等各人文学科领域,已经产生了超越传统研究方法的学术成果。数字人文评价是将一些额外的价值附加到数字人文项目或成果身上,突破传统的以刊评文、以量胜出的单线或线性的评价,从多维和复合的角度为数字人文发展赋能赋值。同时,数字人文在由中级阶段向高级阶段演进的过程中,势必也需要在数字人文评价体系建设的基础上,规范相关研究及成果,引导数字人文的良性发展。由此,数字人文评价开始引起学界的关注,日益成为实践和研究领域的新课题。

数字人文是否需要不同于传统学术的评价体系?基于数字人文研究成果的非典型性实践性,如何对多样性的成果形态进行认证?如何能够公正客观认识跨学科成果?从数字人文发展阶段来看,这些问题的回答有利于推动数字人文发展逐渐走向规范化、科学化^[9]。上文分析了数字人文的三个核心要素和三个发展阶段,对其成果评价也可从这三个核心要素入手

(见表1)。鉴于不同的发展阶段数字人文的评价侧重点不尽相同,同时指标体系需要具有足够弹性,尽可能包容数字人文所有的成果形式,

本文主要是提供一种评价的维度或方向,希望学界同仁共同关注数字人文评价理论研究,构建更完备的指标体系。

表1 数字人文评价的维度及指标

维度	指标	说明
学科性	学科思维体系	主要包括是否突破传统学科思维与体系模式,呈现出新的思维范式、学科特征、理论方法和体系结构等
	学科资料整合	主要包括资料的权威性、珍稀性、完整性、系统性等
	学科数据解读	主要包括数据的合理性、专业性、深刻性、有效性等
	学科交叉创新	主要包括学科的交叉性、融合性、跨界性、创新性等
	学术成果认定	主要包括成果应用的多元性、实用性、开放性、社会性、传播性、反思性,以及合作标注的规范性、协作性、合理性、贡献度等
专业性	数据编制	主要包括资源揭示的广度、深度和粒度,元数据编制的准确性和完整性,数据关联、数据分析等应用的学科匹配度、合理性,人文数据资源建设标准等
	数据管理	主要包括数据存储空间建设,数据管理的科学性和系统性,过程数据的保存完整性、可存储性、可获取性等
	数字实验	主要包括数字实验的科学性、适宜性、可复制推广性、可还原验证性等
	平台建设	主要包括成果展示和应用平台的架构设计、板块设置、功能实现、用户体验等
	教育培训	主要包括面向用户开展的工作坊、座谈、研讨、交流、体验等教育培训形式的多样性,以及平台建设中应用工具、方法的开放性、学习性等
技术性	数字方法	主要包括文本分析(词频、共现、关联、向量、概率)、聚类分类、主题分析、内容挖掘、地理空间分析、社会网络分析等技术
	数字技术	主要包括扫描、拍摄、采样、捕捉、图形设计、3D建模等技术
	技术标准化	主要包括数据标准化、接口标准化、应用标准化等

2.1 学科性:核心要素

学科性是数字人文实践和研究的核​​心要素,也是评价的核心要素。评价一项数字人文成果,最终要看它是否拓展了人文学科的研究领域,是否推动了人文领域的理论创新、知识创新和学术创新,进而促进各人文学科的交叉融合^[10]。对于学科性的评价,即对于学科自身专业领域各个维度的评价,必须要由人文学者深度参与并认可。对于数字人文学科性的评价可从学科思维体系、学科资料整合、学科数据解读、学科交叉创新、学科学术成果等五个方面展开。

(1) 学科思维体系:数字人文是人文学科研

究的一种新的研究范式,突破了传统的思维模式。同时,数字人文又涵盖了传统人文学科的各个方向,例如历史学、哲学、语言学、文学、艺术、考古学、音乐学、文化研究等。近年来,数字人文作为一种新的人文学科研究范式及研究新领域逐步形成了一个新的学术共同体,新学科的特性逐渐显现出来,数字人文的学科体系雏形也逐步显现。虽然距一门独立学科尚有差距,但也体现出一些鲜明的学科特色,数字人文基本概念、思维方式、研究内容等趋同性不断增强。因此,考察是否突破传统学科思维与体系模式,呈现出新的思维范式、学科特征、理论方法或体系结构等,

成为评价数字人文跨学科特性的一项重要指标。

(2) 学科资料整合:这是开展一项数字人文项目研究的基础,其中对资料来源的评价是比较传统的方式,但也是最稳定的方式,主要考察整合的资料是否完整可靠,是否稀缺罕见,重点可以从其权威性、珍稀性、完整性、系统性等角度进行评价。其中,内容准确权威最为重要,它影响到用户是否能获取可靠的知识,尤其是研究者在利用外部数据时,要特别甄别其真实性^[11]。如果一项数字人文项目的资料是独有的、唯一的、罕见的,将其进行数字化整合并向学界或社会开放,其价值就非常大。这是评价数字人文的首要前提,因为无论什么研究项目,不管你用多么先进的技术方法开展,如果基础不稳,那后续的研究都缺乏根基。

(3) 学科数据解读:数字人文项目通过各种工具和方法开展研究后,得到的统计数据或可视化图表需要由学科专业人士进行解释和说明,从而得出研究结论。而这些解释是否具有解决问题的合理性,是否符合学科领域的专业性,是否具备解释学术的深刻性,是否具有指导实践的有效性等,都是开展评价的重要方面。应该考察其解读是否有效推进了学科领域的理论创新和知识创新,取得了传统人文研究方法无法取得的成果。

(4) 学科交叉创新:目前,我国“新文科”建设是为了对接新一轮科技革命所带来的学科发展和交融要求,在“应变”和“求变”中实现哲学社会科学教育及知识生产模式深刻而全方位的变革^[12]。数字人文的跨学科性则成为新文科发展的新生力量,可以促进交叉学科的形成和发展。因此,从学科的交叉性、融合性、跨界性、创新性等角度去评价一项数字人文研究的创新度是符合新文科发展要求的,也是开展其建设的题中应有之义。

(5) 学术成果认定:跨学科是数字人文的最本质属性,学术成果认定也与传统学术成果的认定不同。数字人文项目本身具有广泛的合作性,学术成果中作者概念和传统的已有较大不

同,尤其是不同领域之间的合作。同时,很多数字人文成果不只是论文,还包括虚拟仿真模型、程序、代码设计、数据可视化、数字期刊、数字博客等其他类型。因此,学术合作标注以及学科成果应用显得尤为重要,成为数字人文重要的学术规范之一。在合作标注方面,首先,合作者共同的价值认同和规则认同,以及合作者的协作性和合理性,是数字人文项目得以顺利开展的前提;其次,知识产权的归属,以及合作形式等方面的标注,也是学术规范的表现形式。在学科成果应用方面,当前的数字人文项目更多的是以实践应用型为导向,因此对其成果的评价应该坚持多元性和实用性的原则^[13]。同时,建成的数据库或平台向特定人群或社会公众提供的使用权限,即其开放性、社会性和传播性也是重要的评价指标。此外,数字人文研究还应该反思自身研究的数据基础设施、采用的方法和步骤,以及这些所带来的局限性,或者说未来可以修正的空间^[14]。当然,成果的应用前景、社会效益和价值创造等因素也是必不可少的指标。

2.2 专业性:关键要素

数字人文虽然具有鲜明的跨学科性,但也具有非常鲜明的专业性,需要具有一定深度的专业知识,不是所有人都能轻易入手从事数字人文工作。而图情档学科利用其信息组织、检索、分析等专业知识为数字人文提供了方法论指导,也是连接学科和技术的桥梁和纽带。数字人文的专业性评价主要体现在数据编制、数据管理、数字实验、平台建设、教育培训等方面。

(1) 数据编制:数据是数字人文的基础与核心,因此数字人文的数据设计是关键一环,即数据的编制,数字人文中的数据编制包含元数据定义、数据建设、数据关联的建立以及人文数据资源建设标准化等,这些都需要一定的专业性知识。根据数字人文的研究需要,对资源的广度、深度和粒度进行多维度揭示,并编制准确和完整的元数据,将不同维度的数据建立关联,通过一

定的标准最终完成人文数据建设。因此,透过数据编制揭示数据之间的关联和指向,透过时间和空间的维度呈现出不同研究视角,以立体、多元、多维的方式展现数据之间的关联,应该就是数字技术突破传统人文研究的最大助力。

(2)数据管理:数字人文研究过程中会产生大量的过程数据,其作为重要的研究过程的体现,应该予以保存和提供给后续研究者继续利用。如何对过程数据进行存储和管理涉及一些专业性知识,数据存储空间建设的情况、数据管理的科学性和系统性,以及数据保存的完整性和可获取性是评价的主要方面。当前已有很多学术期刊要求数字人文研究在提交论文的同时提供原始数据,正是对这一评价维度的最好体现。

(3)数字实验:数字人文项目往往具有实验属性,很多高校也建设了数字人文实验室。一方面需要对数字人文实验设计的科学性、学科适宜性等开展评价,另一方面从实验本身的特性出发,其应该是具有可复制推广性、可还原验证性的,因此需要检验这样的实验是否可以被其他研究者再次实施和验证,是否具有反复进行的可能性。

(4)平台建设:数字人文研究成果中的数据库平台应当成为评价的重要内容,正如 CBDB 成为数字人文的一项标志性产品或成果,可见应用型的实际产品更应受到重视。对成果展示和应用平台的评价可以借鉴传统的对数据库评价的维度,从其架构设计、板块设置、功能实现、用户体验等角度进行评价,此外需重点关注其实际利用率、社会影响力,以及产生的实际效益和社会价值。

(5)教育培训:从目前数字人文相关教育与培训实践来看,参与者可能更多是图情档领域的馆员、学生、学者和实践者,而人文学者的参与面还不是很大。此指标一方面可以促使数字人文项目采取更多的教育培训方式,让更多参与者了解数字人文项目和成果,扩大数字人文理论与技术的认知度和参与面;另一方面使项目从人文学者本身的学术训练视角出发,更多地将视角关注

到人文学者身上,真正思考如何让人文学者掌握数字技术的相关工具与方法,带动人文学者一起深入参与数字人文研究,而这才是教育培训的真正志向所在。例如,可以通过构建数字人文数据中台、工具中台和方法中台等,更好地嵌入跨学科支撑中,让更多的人文学者能够有机会训练和使用这些工具与方法。

2.3 技术性:必备要素

数字信息技术对人文学科研究的资源类型和研究方法产生了巨大影响,数字人文正是数字技术与人文学科交叉而形成的跨学科研究领域^[15]。信息技术、自然语言处理技术、语义网技术、文本挖掘技术等应用于人文数据分析中,助力人文领域的学术创新和知识创新。数字人文技术性评价主要体现在数字方法、数字技术以及技术标准化等方面。

(1)数字方法:对人文研究的探索只有依靠现代技术与数字人文工具及方法,挖掘隐藏在文本内外的各种知识、关联以及透过文本表达出来的情感因素,才可以突破传统人文研究的局限。当前数字人文领域应用较多的技术主要涉及文本分析(词频、共现、关联、向量、概率)、聚类分类、主题分析、内容挖掘、地理空间分析、社会网络分析等^[16],对这些技术是否合理选择和科学利用是评价的重要内容,主要看其是否符合研究的实际需求,而非为用而用,生搬硬套。

(2)数字技术:数字技术作用于传统人文能使其实现动态和知识关联,标志着“数据+算法+算力”正在成为人类应对不确定性的主要手段之一,传统文科必须拥抱数字技术,以提升自身研究的准确性与科学性^[17]。当前数字人文主要运用到的数字加工技术包括扫描、拍摄、采样、捕捉、图形设计、3D建模等,可以从对这些技术选用的针对性、合理性、有效性,以及技术运用的专业度和熟练度等角度来开展评价。

(3)技术标准化:数字人文的技术多元且异质,在研究中需要尽可能地统一和规范。

因此,需要对数据标准化、接口标准化、应用标准化等的实践予以评价,在评价中不断探索和实现技术标准化。

3 图情档在数字人文发展和评价中的作用

3.1 图情档利用其专业性推动数字人文发展

图情档应利用其专业优势,推动数字人文的发展和 innovation。第一,图情档等机构和学科继续为数字人文研究提供数据基础设施,为人文社科数据分析提供思想、理论和方法。第二,发掘图情档机构中人文社科人才的潜力,利用资源和数据优势,使本领域的跨学科人才真正融入人文社科的研究之中,产出图情档的人文研究新思想和新理论。第三,与高校和科研院所的人文社科机构深度合作,共同推动人文领域的思想创新和理论创新,真正让人文社科领域的专家学者认可数字人文。第四,建设数字人文实验室等场所为高校跨学科数字人文研究提供各类基础设施、交流协同空间和技术方法指导,并通过数据中台等理念开展数字人文服务与支持。

通过已有研究成果可以发现,数字人文研究呈现出明显的跨学科特征,涵盖人文社会科学领域和自然科学领域的相关学科。除传统人文学科之外,图情档作为数字人文这个新型跨学科中的重要成员,在经过知识转型、重新定位之后以多重身份参与数字人文相关研究。首先是扮演“应用者”角色,图情档借助数字技术,在资源建设、知识服务、空间设计等方面更“聪明”、更“智慧”^[18];其次是扮演“合作者”“协调者”和“支持者”的角色,数据是数字人文研究最基础的保障,而图情档领域以其先天存储信息资源的优势能为人文学者提供从结构化数据到非结构化数据的挖掘机会^[5]。

对于图情档领域研究者而言,无论是从应用者的角度还是从合作者、支持者的角度,都需

要紧密对标新文科建设的方向,在了解人文学者研究成果的基础上,加强与人文学者、计算机领域相关人员的紧密合作。一方面可以基于数字人文项目,加强与人文学者的沟通,准确把握人文学者的信息需求与研究思维,从人文研究的视角开发更多的数字工具和方法;另一方面,通过数字人文项目,吸引更多的人文学者尝试通过合作,认识到数字人文研究的无限魅力,这不仅可以通过合作产出更多有价值的成果,促进数字人文向广度和深度延展,而且能够以点带面进一步突显数字人文的跨学科特性,为数字人文发展增添活力。

3.2 图情档利用其专业性促进数字人文评价

数字人文评价本身是专业性较强的一项工作,无论是指标体系构建,还是评价实践本身,都需要专业性的知识、专业性的人员和专业性的组织。

(1)专业性知识优势:学术评价体系研究和构建是图情档的专业特长,也在实践中积累了丰富的经验。图情档领域的研究者具备设计和编制相应评价指标体系的知识优势,熟悉元数据的标引、关联数据的整合、科研数据的管理,以及各类数字人文工具和方法的应用、数据库平台的构建等关键领域的知识,具有开展数字人文评价的天然条件。应用这些专业性知识和技能,对于构建相对健全的数字人文评价体系并不断修订完善,具有很强的指导意义。

(2)专业性人员优势:针对上述专业性知识,图情档领域配置有“术业有专攻”的各类专业性人员,包括熟悉图情档业务的人文学科背景的人员,以及熟悉图情档业务和需求的专业技术人员。一方面,数字人文评价离不开既熟悉元数据标引和相应工具方法,同时又兼具一定人文学科背景的人才,他们可以在评价中起到联系人文学者与技术人员的“桥梁”作用,并对人文学者开展数字人文研究所应用的技术方法等进行评价;另一方面也离不开专业的技术团队的支持与保障,他们不仅为数字人文项目的顺利开展提供强有

力的技术工具和方法指导,也能对其中的技术难题攻关和技术问题解决予以专业性评判。

(3)专业性组织优势:数字人文评价目前尚未形成较为完善的工作机制,缺乏统一的组织和规范的流程,具体请哪些专家来评价也存在争议。在这一问题上,图情档机构可以发挥其中介“桥梁”作用和组织功能,为来自人文学科领域、图情档领域和专业技术领域的各方专家搭建一个数字人文评价委员会平台,组织相应的专家应用其开发的评价指标体系来协同评价,并在实践过程中不断完善评价指标体系,进一步搭建合理的组织平台。

4 小结

数字人文研究以跨学科为特点,以项目为单位,以合作为主体,以数据为基础,以技术为手段,以平台为支撑,与传统的人文社会科学研究在对象、范式和成果性质方面都不尽相同,需要重构学术评价理论和评价指标体系^[13]。结合数字人文研究特点及发展阶段,从学科性、专业性和技术性三个维度探讨数字人文评价,可以看出:学科性是核心要素,评价数字人文成果最

终要看它是否推动了人文领域的实践创新和学术创新;专业性是关键要素,图情档学科和机构为数字人文建设数据基础设施,并利用其信息组织、检索、分析等专业知识提供方法论指导,也是连接学科和技术的媒介和纽带;技术性是必备要素,信息技术、自然语言处理技术、语义网技术、文本挖掘技术等应用于人文数据分析中,助推人文领域的学术创新和知识创新。因此,我国数字人文研究需要瞄准“新文科”建设目标,突破以图情档为主体的研究局面,加强跨学科交流与合作,在更大范围内引起人文学者乃至其他领域学者的重视,从而进一步推动人文社会科学和新型技术融合,培育良好的数字人文生态,以促进各领域研究人员进行跨学科、跨机构、跨地区甚至全球性的数字人文研究和实践^[19]。图情档领域需要承担起数字人文发展的使命和责任,积极利用数字人文研究的契机,发挥专业知识优势、技术应用优势、专业人才优势和专业组织优势,加强跨学科交流与合作,协助推动人文学科领域的理论创新、知识创新、学科创新和方法创新;同时,借鉴已有评价研究成果和评价实践经验,引领开展数字人文成果的评价。

参考文献

- [1] Berry D M. Understanding digital humanities[M]. New York: Palgrave Macmillan Press, 2012.
- [2] 崔凤娇. 探索数字人文[J]. 跨文化研究, 2019(2): 210-276. (Cui Fengjiao. Exploring digital humanities[J]. Cross Cultural Studies, 2019(2): 210-276.)
- [3] 许苗苗, 邵波. 我国数字人文发展的脉络、问题及启示[J]. 图书馆学研究, 2020(14): 2-10. (Xu Miaomiao, Shao Bo. Context, problems and enlightenment of digital humanities development in China[J]. Library Science Research, 2020(14): 2-10.)
- [4] 周文杰. 知识资源的序化与数字人文的结构化知识资源基础——基于知识地图和认知结构学说的解析[J]. 情报资料工作, 2020(6): 79-87. (Zhou Wenjie. The ordering of knowledge resources and the basis of structured knowledge resources for digital humanities: an analysis based on knowledge map and cognitive structure theory[J]. Information and Data Work, 2020(6): 79-87.)
- [5] 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(1): 17-34. (Zeng Lei, Wang Xiaoguang, Fan Wei. Smart data in the field of library archives and its role in digital humanities research[J]. Journal of Library Science in China, 2018(1): 17-34.)
- [6] 大卫·M. 贝里, 安德斯·费格约德. 数字人文: 数字时代的知识与批判[M]. 王晓光, 等, 译. 大连: 东北财经

- 大学出版社,2019。(Berry D M, Fagerjord A. Digital humanities: knowledge and criticism in the digital age [M]. Wang Xiaoguang, et al, trans. Dalian: Northeast University of Finance and Economics Press, 2019.)
- [7] 项蕾,许婷,谭天,等.数字与文学的对话——“数字人文规范对传统文学研究方法的挑战”研讨会纪要[J].中国现代文学研究丛刊,2020(8):158-176。(Xiang Lei, Xu Ting, Tan Tian, et al. Dialogue between numbers and literature: summary of the seminar on “challenges of digital humanistic norms to traditional literature research methods” [J]. Chinese Modern Literature Research Series, 2020 (8): 158-176.)
- [8] Duro A. Humanities computing activities in Italy[J]. Computers and the Humanities, 1968,3(1):49-52.
- [9] 杜娟,雪映.数字人文:何以评价[N].社会科学报,2020-10-08(004)。(Du Juan, Xue Ying. Digital humanities: how to evaluate [N]. Social Sciences, 2020-10-08 (004).)
- [10] 司莉,陈辰,郭思成.中国图书馆学的应用实践创新及发展研究[J].中国图书馆学报,2021(3):23-42。(Si Li, Chen Chen, Guo Sicheng. Research on application, practice, innovation and development of library science in China [J]. Journal of Library Science in China, 2021 (3): 23-42.)
- [11] 赵翔翔,刘周颖,刘炜,等.创意类开放数据竞赛作品评价指标体系构建与测定——以数字人文项目为例[J].中国图书馆学报,2020(2):75-95。(Zhao Yuxiang, Liu Zhouying, Liu Wei, et al. Construction and measurement of evaluation index system of creative open data competition works: taking digital humanities project as an example [J]. Journal of Library Science in China, 2020 (2): 75-95.)
- [12] 马费成,李志元.新文科背景下我国图书情报学科的发展前景[J].中国图书馆学报,2020(6):4-15。(Ma Feicheng, Li Zhiyuan. The development prospect of library and information science in China under the background of new liberal arts [J]. Journal of Library Science in China, 2020(6):4-15.)
- [13] Bradley A J, EL-Assady M, Coles K, et al. Visualization and the digital humanities[J]. IEEE Computer Graphics and Applications, 2018(6):26-38.
- [14] 王涛,陈静,等.“数字人文的学术评价体系:定义与规范建构”研讨会发言汇编[J].数字人文,2021(1):1-20。(Wang Tao, Chen Jing, et al. Symposium on academic evaluation system of digital humanities: definition and normative construction [J]. Digital Humanities, 2021 (1): 1-20.)
- [15] 欧阳剑.大数据视域下人文学科的数字人文研究[J].图书馆杂志,2018(10):61-69。(Ouyang Jian. Digital humanities research in humanities from the perspective of big data [J]. Library Journal, 2018 (10): 61-69.)
- [16] 单蓉蓉,陈涛,刘炜,等.国际图像互操作框架及拓展应用[J].图书馆杂志,2021(5):89-95,123。(Shan Rongrong, Chen Tao, Liu Wei, et al. International image interoperability framework and extended application [J]. Library Journal, 2021 (5): 89-95,123.)
- [17] 王丽华,刘炜,刘圣婴.数字人文的理论化趋势前瞻[J].中国图书馆学报,2020(3):17-23。(Wang Lihua, Liu Wei, Liu Shengying. The theoretical trend of digital humanities [J]. Journal of Library Science in China, 2020(3):17-23.)
- [18] 赵洪波,罗玲,李大莉.数字人文视域下智慧图书馆建设的模式与路径[J].图书馆学刊,2019(9):110-113,122。(Zhao Hongbo, Luo Ling, Li Dali. The mode and path of smart library construction from the perspective of digital humanities [J]. Library Science Journal, 2019(9):110-113,122.)
- [19] 宋玲玲,郭晶晶.科学知识图谱视角下国内外数字人文领域研究分析[J].图书馆杂志,2020(7):26-36。(Song Lingling, Guo Jingjing. Research and analysis of digital humanities at home and abroad from the perspective of map of scientific knowledge [J]. Library Journal, 2020 (7): 26-36.)

蔡迎春 上海外国语大学图书馆常务副馆长,研究馆员。上海 200083。

(收稿日期:2021-02-22;修回日期:2021-06-19)