

搭建历史研究共享平台的探讨

——以口述历史为切入点

李二苓

(中国社会科学院当代中国研究所,北京 100009)

摘要: 作为史料,口述历史的真实性受到质疑。较其他史料而言,口述历史的采集、编辑、呈现等需要更加规范,以确保其真实性。口述历史具有跨学科以及连接学术界与大众的媒介特点。因此,以口述历史作为构建未来历史研究共享平台的切入点,有利于满足比较研究、量化研究、数字学术、公共史学、跨学科研究等领域对“数字人文”提出的需求,不仅丰富了资料获取方式,并可通过数据分析发现新的知识。同时,打造一个资源、知识关联、知识服务均可持续扩展的公共文化交流平台,可使史学从基础学科向应用延伸,强化知识普及、文化引领等社会服务功能。

关键词: 口述历史;立体资料数据库;历史研究共享平台;社会服务;新业态

Doi: 10.13806/j.cnki.issn1008-7095.2022.0.009

20世纪80年代以来,随着社会史、文化史的兴起,我国学术界的口述历史记录和研究日益丰富,口述历史数据库建设也随之起步。目前国内关于口述历史的研究大多从两个方面切入,一是如何做口述历史,即口述访谈具体的操作步骤和注意事项;二是如何做口述历史研究,即运用访谈材料进行学术研究,以论文或著作的形式考辨历史事实。目前,有很多口述历史项目及数据库的建设,但其经验探讨仍集中于口述历史采集方法,对于怎样呈现、用好口述历史资料,传播口述历史研究成果等还缺乏理论思考。在原本已经浩如烟海的现当代史料中,再增添回忆录等新史料,研究者有限之精力难以处理这般无垠之史料。随着“数字人文”的发展,在新一轮的“史学革命”中,历史资料如何数字化呈现以便服务科研的问题亟待解决,这一问题的解决或许会带来史学研究的新发展。

数据库主要是资料的集聚,数据平台则更强调交互功能。在数据库技术发展日新月异的背景下,历史资料的获取、呈现、服务科研等或也可随之升级,历史类数据库应向历史研究共享平台转变。

随着口述历史专业化的发展,研究者对口述资料的质疑和批判与日俱增。尤其是口述历史的真实与否,成为其能否作为史料的关键。因而,口述历史资料的编辑加工和呈现方式需要更加规范,可以为平台运营各类史料树立典范。相较于其他史料的搜集、整理和运用,口述历史更具有跨学科和大众化的特点,受科技的影响更大,更有可能成为构建历史研究共享平台的切入点和突破口,有利于改善专业历史学与社会大众的疏离状态。据此,本文希图以口述历史对数字化网络平台的需求为切入点,为建设历史研究共享平台张本。

一、口述历史的特点与立体资料数据库

打破各自为政的信息壁垒,减少资源的重复投入,建设资料类型丰富、资源互相关联的立体

作者简介:李二苓,中国社会科学院当代中国研究所助理研究员。

资料数据库,是搭建历史研究共享平台的基础。

(一) 口述历史之初心——补史料之不足

国际上,大多数学者认为1948年美国哥伦比亚大学的历史学家艾伦·内文斯(Allan Nevins)着手进行美国重要人物的访谈工作是口述历史建立的标志。其做口述历史的初心是补足重大历史事件的史料,主要访谈对象是没有自传的领袖。

这一新研究路径对中国学术界的影响,首先体现在台湾地区。1955年,郭廷以在创办“中央研究院”近代史研究所时,就尤为重视口述历史。20世纪60年代该所与哥伦比亚大学东亚研究所合作,并获得美国福特基金会的资助,对北洋政府、国民政府的将官展开口述历史访谈。1984年,吕实强所长成立“口述历史组”,开展对台湾地区各界代表人士的访谈,以及对二二八事件等政治案件的专题访谈。这些研究成果主要通过《口述历史》杂志发表和“口述历史丛书”出版。从内容上看,台湾地区的口述历史局限于重要历史人物和重要历史事件。

中国大陆向来有口述传统,新中国成立初期就开始了政协文史资料的搜集。毛泽东号召撰写“四史”(家史、厂史、社史、村史)等也是以访谈、田野的方式来搜集历史资料。不过,现代学术层面的口述历史,则是从20世纪80年代逐渐发展起来。由于口述历史受限于当事人健在这一客观条件的制约,现当代史是口述历史应用的主要场域。重要历史人物和重大历史事件仍然是口述历史的重点。特别是在新中国史领域,有“当代中国口述史”系列丛书等一批代表性的成果陆续出版。^①另外,受到20世纪60年代以来欧洲学术界与左翼政治运动结盟的社会、文化历史研究^②和20世纪70年代美国史学“自下而上”方法的影响,那些被忽视的普通人的生活也被纳入口述历史研究范围,抗战老兵、少数民族、妇女、知青、非物质文化遗产传承人等,都成为口述历史的访谈对象。口述历史成为中国当代社会史的重要基石和源泉。^③

(二) “不可靠记忆”与比较研究

口述历史是基于访谈对象回忆而形成的,具有一定的不可靠性和主观性。不可否认,主观性与不可靠性同样存在于其他史料中,即便是档案资料,有些内容也不见得就是历史真实。董国强指出目前历史学界普遍存在的问题是过度依赖档案资料,即所谓盲目的“档案拜物教”倾向,导致研究成果缺乏创新性,也很难听到普通民众的声音,达不到反映多元社会关系结构的目的。^④口述历史等获取历史信息的新方法对于现当代史研究来说更重要。运用史料是否能够得到客观的研究成果,关键取决于历史研究者的功底。

比较研究是去伪存真的有效方法之一。将不同社会阶层、不同视角的人所留下的史料对比阅读,才能全面、深刻地反映历史样貌。研究者只有充分了解已有研究,对勘不同来源的资料,才能做出经得起推敲的学术成果。

数据库的首要特征是资料浩繁。档案、报刊等核心史料是历史数据库必不可少的收录资料,在此基础上增加回忆录、民间文书、日记、书信等原始资料,以及著作、论文、报告等研究成果,形成立体的资料体系。数据库中既有从上而下,又有从下而上的观察;既有当事人,又有旁观者的看法;既有失败者,又有胜利者的回忆;既有当时的记载,又有后人的研究;除了政治、经济之外,还包括日常生活、音乐、宗教等主题;在搜集访谈音、视频的同时,也征集与访谈相关的

① 姚力.再论国史研究与口述历史[J].中国科技史杂志,2009,30(3):284-285.

② 持历史书写应关注底层历史的观点的代表性人物是英国史学家保罗·汤普森,他1978年出版了巨著《过去的声音——口述史》。保尔·汤普森.过去的声音——口述史[M].覃方明,渠东,张旅平,译.沈阳:辽宁教育出版社,2000.法国年鉴学派也提出“the history from bottom up”(由下向上观察历史)的观点。

③ 姚力.试论口述历史对中国当代社会史研究的启示[J].当代中国史研究,2012,19(4):33-39.

④ 董国强.中国当代史研究方法论两题[J].中共党史研究,2021(1):152-155.

信件、日记、照片、回忆录和受访者捐赠的其他资料,丰富数据库资源类型,立体地展现历史。

任何有价值的历史研究成果都不是运用单一史料、从单一的角度来分析的,而是需要“上下左右”求索与论证,充分运用文献学、校勘学等史学研究方法,以求得客观公正的史家评判。立体资料数据库是历史研究对史料数据库的要求。

(三) 口述历史的“文化转向”与生命史全流程发布

历史学对史料的解读不应停留在史料之所指,而应考虑史料背后的生成机制,掌握史料产生的整个过程。对口述历史而言,即不宜仅访谈历史大事件,而应做生命史。生命史全流程发布是口述历史对立体资料数据库所提出的要求。

20世纪70年代末,西方学术界针对口述历史的真实性展开了一系列理论探讨。受后现代理论等的影响,口述历史成为连接历史与当下的桥梁,记忆成为历史分析的对象,形成了所谓的口述历史的“文化转向”(cultural exchange)^⑤。保罗·汤普森曾指出:“社会统计资料并不比报纸报道、私人信件或者出版的自传更多地代表绝对真实。像记录下来的访谈资料一样,它们要么从个人的立场要么从集合体的立场,全部代表着对事实的社会感知;并且所接受到的东西是社会意义,并且正是它必须予以评介。”^⑥口述历史的受访者所说的记忆,“它不一定是过去发生的事实,却反映了个人的认同、行为、记忆与社会结构间的关系”^⑦。口述历史工作者不仅要保存大量访谈和观察资料,还“必须注意到人们记得什么,他们为什么会忘记或为什么会过去的对过去的事件重新整理”^⑧。

因此,口述历史要做生命史,要“去了解他的全部,然后再进入我们感兴趣的那个地方”。只有对受访者的生命历程的前因后果有充足的了解,才能分析受访者是怎么想的,为什么要那么说。

随着电子技术和网络技术的发展,口述历史每个阶段的成果物均可以上传到网络,低成本本地大规模发布。因此,口述历史访谈的公共呈现已经成为口述历史实践中最具活力的内容之一,其公共呈现的完整与否也一定程度上决定了口述历史能否做到客观。

口述历史的特点之一即“双重主体”性,一个主体是作为研究者的访谈者;另一个主体是作为历史当事人的受访者。后者既是研究对象的客体,又作为旁观者、参与者成为历史研究的主体。^⑨口述历史是访谈者和受访者共同创造的过程^⑩,因此发布访谈过程中有授权的全部资料有助于区分访谈者与受访者各自的表达,更准确地还原事实,充分展现身份认同、价值观、隐喻等文化与社会的信息。西方口述历史学界甚至会讨论读抄本与听录音的差别。^⑪究竟声音和影像会不会干扰研究者对受访者口述的客观解读尚待讨论,但这些探讨进一步强调了口述历史全流程发布是确保其史料价值的前提。^⑫

⑤ Thomson A. Fifty Years On: An International Perspective on Oral History[J]. The Journal of American History, 1998, 85(2): 586.

⑥ 保尔·汤普森.过去的声音——口述史[M].覃方明,渠东,张旅平,译.沈阳:辽宁教育出版社,2000:131.

⑦ 定宜庄.关于《最后的记忆》一书的说明[J].北京航空航天大学学报(社会科学版),1999,12(23):47.

⑧ 唐纳德·里奇.口述史的发展历程[M]//唐纳德·里奇.牛津口述史手册.宋平明,左玉河,译.北京:人民出版社,2016:15.

⑨ 左玉河.口述历史与国史研究[J].当代中国史研究,2016(3):48-49.

⑩ 罗纳德·格里,柯里·罗比,玛丽·克拉克.口述历史与档案馆:以哥伦比亚大学口述历史中心的经验为例[G]//杨祥银.口述史研究:第一辑.北京:社会科学文献出版社,2014:28-39.

⑪ Cohen S. Shifting Questions: New Paradigms for Oral History in a Digital World[J]. The Oral History Review, 2013, 40(1): 154-167.

⑫ 为了解决口述历史的真实性问题,有些机构采取出版口述历史整理稿的方法以求获得权威性,但却难以展现口述历史的生动性和双重主体性。例如南京民间抗日战争博物馆的抗战老兵访谈,先将访谈成果发表于地方文史资料,再在博物馆官网上展示文史资料的登载内容。

数据库发布成本较低,利于大批量发布。因此,数据库不仅要发布整理出来的回忆录、综述片,充分体现口述历史的可读性、艺术性,以便吸引读者,而且,访谈的提问大纲、音视频原始文件和抄本也有必要发布,以便学术研究者追根溯源。只是在抄本中需要把有史实错误的地方标注出来,并注明出处,以免误导读者。^⑬

(四)“大数据”普遍关联的立体资料数据库

立体资料数据库主要包含两层含义。一是资料类型、来源的多样化,资料生成过程的完整性,以满足“上下左右”论证的需要。建设数据共享开放平台是学界广泛呼吁并持续努力推进的,例如“抗日战争与近代中日关系文献数据平台”通过广泛搜集史料、免费使用,达到打破史料壁垒、突破史观局限、横向纵向史料互证的目的。^⑭但是,目前“数据孤岛”、重复建设现象仍较严重,大多数数据库仅限于原创机构内部使用,而一些开放的营利性数据库,则主要局限于图书、期刊等资料类型。造成这些现象的主要原因是版权壁垒,特别是现当代的资料,档案可能暂未解密,图书、报刊、实物资料、口述资料等又各有所属。因而,需要在整合图书馆、档案馆、博物馆、出版单位、研究机构等方面有所突破。

二是资料链接立体化,实现知识关联,方便追根溯源。数据库不仅要有精准检索及二次筛选、高级检索等基础功能,更要实现知识的关联以及知识与资料的关联。(1)全文数字化是增加功能的基础。相较于图片扫描,全文数字化成本较高,因此有些大型公益型数据库(如“抗日战争与近代中日关系文献数据平台”)往往做不到全文数字化。(2)信息整理的规范化是保证资料学术价值和资源共享的基础。要充分尊重资源原有的形成环境和保存形式,形成资源整理、上传的国家级规范。(3)知识关联功能需同时满足两个层面的要求才能对史学研究有切实的帮助:一是知识与知识的关联,指要将每一个历史知识与其他相关的历史知识关联。二是知识与资料关联,知其然亦要知其所以然,指每一条知识信息或知识关联信息都要有据可依,方便读者快速查阅信息来源资料。目前,在历史学界影响较大的人物资料库——“中国历代人物传记资料库”(CBDB)利用地理信息系统,将亲属关系、社会关系等作为勾连人物知识的脉络。但是目前“中国历代人物传记资料库”被认为是好评如潮却少有问津。好评如潮的原因是该库抓住了历史研究者的痛点,做到了知识之间的关联。少有问津的主要原因在于该库使用过于复杂,知识关联与史料脱节,从人物界面无法直接获取相关史料。数据分析与文本阅读的脱节也是目前人文学界诟病数字人文的重要原因。(4)文献比对功能。该功能是指将涉及同一历史事物的文字进行比对,展示不同资料对同一事物的不同叙述,方便读者去伪存真、分析流变。其技术难点在于历代用语各有不同,自然语言处理的难度较大。目前能够实现同主题文献比对功能的数据库,有人民出版社的“人民金典语义查询系统”,比对范围限于共产主义经典著作。

二、发现新知识、共享科研过程的历史研究平台

现在有一种提法叫“量化数据库”,这一概念不仅强调资料的数量巨大,还强调“可适用于统计分析软件”进行量化研究,因此它带来的科研革命不仅是“资料革命”,更是研究方法的革命。在利用数据库做量化研究时,文本分析、社会网络、时空地图等方法运用得越来越多。新技术“真正介入研究的核心”,使学者们“从文献叙述中发现问题转变为从数据分析中发现问题”^⑮。历史研究平台在促成研究方法的革命上则要更进一步,不仅要满足学术研究的功能性要求,更

^⑬ 目前有学者主张在抄本中直接修改违背史实的内容。笔者不太赞同这一做法,建议保留受访者原意,将考证史实的结果标示在旁,并注明所谓史实之出处,读者才能明确何为受访者的记忆,何为编辑加工者的观点。

^⑭ 罗敏.“抗战文献数据平台”与创新中共抗战史研究的方向和可能[J].中共党史研究,2020(6):157-159.

^⑮ 梁晨.量化数据库:“数字人文”推动历史研究之关键[J].江海学刊,2017(2):162-164.

要突出交互性和互通性。

(一) 数字学术与知识服务

数字人文在学术界的发展存在一个从量化研究到数字学术的转型。20世纪中期以来,学界大力提倡跨学科研究方法,于是统计学等数量分析方法在历史学、社会学、人类学、哲学、文学等领域广泛应用,定性研究与定量研究密切结合,使得人文学科进一步精确化,如美国的“新经济史”开创了计量史学之先河。20世纪80年代以来,计量史学扩展到中国史学的各个领域,成为一种潮流。随着计算机技术和互联网技术的发展,知识的获取、分析集成、展示和传播的体系发生了变化,“大数据”一词孕育而生,学术研究不再限于计量,学者们希望借助技术手段实现多维的知识关联。数据分析不仅仅是为研究提供论据,更可以发现新的知识。以大数据为支撑,通过技术手段,用数据发现新知识的学术研究,可以称为“数字学术”。就史学而言,学界则称为“数字史学”。图书馆界对这一发展现状的关注度最高,学术界对数字人文评价不一。^{⑯⑰}

目前,基于数字学术而建设的结构性数据库往往缺乏共享机制,导致在信息搬运、组织标引等方面造成重复劳动的现象。其原因大致是学术界注重通过著作、论文展现数据库功能和研究成果,而缺乏知识服务意识。所谓知识服务,即以领域知识体系建设为基础,融合相关技术,构建领域知识汇聚机制,对领域知识进行采集、沉淀、组织,设计并生产知识产品,以知识产品为纽带与连接器,连接知识创造与知识应用,提供多样化服务模式,促进知识创新与再创造,支持各领域生产活动、学术研究、包括政府部门在内的各类机构的决策,满足社会公众以知识为核心的文化消费。知识服务的目的是向用户提供可以解决问题的知识资源。^⑱在信息爆炸的数字时代,信息精准化的智能知识服务已成趋势,只有基于用户即时互动的平台,所生产的知识才是最有用的知识,所提供的服务才是最有效的服务。

数字学术在建构结构性数据库时,应加强知识服务意识,向共享、互通方向发展。如果不以互通、共享为建库目标,那么就很难有效整合工作流程烦琐的需要耗费较大人力、财力和时间成本的基础工作,很难推动数字学术的持续、快速发展。中国人民大学“数字清史实验室”^⑲、香港科技大学李中清-康文林研究团队在构建数据库时,关注到了自建多个数据库的匹配连接,并共同致力于数据开放^⑳。这些探索将有助于数字学术研究成果的积累和科研过程的共享。

(二) 元数据多样化与编辑后台前台化

重大历史事件和重要历史人物不再是研究的全部,量化历史研究也早已从经济史延展到社会史、文化史,开拓新领域和学术深耕越来越成为研究的发展趋势。面对如此多样、复杂的学术研究,如何精准满足学者需求,提供更有效的知识服务,成为越来越难做的工作。

万变不离其宗。数字学术崛起,但并非每位学者都具备统计学等学科的知识 and 计算机技能,因此需要借助软件工具处理海量的资料。在搭建数据库时,时间、地点、人物、重大事件、学科分类、文献类型、专业词汇等仍是最基本的元数据,但是随着使用需求的多元化,让用户参与元数据的制定及标引,将是学术平台的发展方向,使“人文学研究者不再是被动选择既有的数据库,而是参与数据库建设过程,由其自身研究需要引导数据库开发,数据库开发过程也就成为其研究的一部分”^㉑。

^⑯ 包伟民,李剑鸣,王日根,等.“大数据时代史学研究的理论与方法”笔谈[J].史学月刊,2018(9):5-7.

^⑰ 邱伟云,严程.数字方法与传统人文更需迈向多元共生关系[N].社会科学报,2021-02-04(5).

^⑱ 刘广峰,张涛,孙金,等.专业出版知识服务标准体系构建研究[J].出版发行研究,2018(12):49-51.

^⑲ 夏明方.大数据与生态史:中国灾害史料整理与数据库建设[J].清史研究,2015(2):67-82.

^⑳ 梁晨.用大规模量化历史数据库检验中国的长期代际遗传[J].南京大学学报(哲学·人文科学·社会科学),2019(2):91-96.

^㉑ 赵思渊.地方历史文献的数字化、数据化与文本挖掘:以《中国地方历史文献数据库》为例[J].清史研究,2016(4):26-35.

并非每位学者都需要参与数据库建设项目,而是可以通过将数据库著录标引的编辑后台提供给用户使用的方式,达到每位学者都加入数据库建设的目的。借此,平台也能实现研究过程的共享,形成在研究中不断生长、互惠共赢的发展模式。方便的量化分析工具也可以使学者从烦琐的量化研究中跳脱出来,做到量化与质性研究的兼顾,凸显人文研究的主体性。

目前已有一些数据库在文本分析、数据处理功能上有所推进,有些还获得了不错的销售业绩。例如,“台湾历史数位图书馆”(THDL)提供词频分析、契约文书买卖角色分析、人物相关性分析、契约空间分布展示等功能;“盖尔学者数据库”(Gale Scholar)整合了国外常用的第三方数据挖掘和数据分析软件,为用户提供词库、词频分析工具和可视化展示。

(三) 从口述历史需求出发的精准推送功能

口述历史更需要知识服务的加持。一方面,网络的发展使得任何人在任何地方都能参与到口述历史项目中来,口述历史的采集量与日俱增,需要引导用户快速、有效地获得目标信息。另一方面,口述历史资料最为庞杂、分散,更需要交互性的平台功能,聚焦热点和前沿问题。

就目前的出版而言,纸质书的作者较难了解读者购买后的阅读情况,作者与读者的互动性很弱。大多商用的学术性专题数据库,运营后台往往是以内容资源为根本,用户行为分析功能薄弱,难以全面把握用户的网上行为基本数据,难以分析用户的来源路径、兴趣点、知识关联、学术产出。如何向用户推送数据库中个人感兴趣的优质资料?如何挖掘出用户普遍关心的问题从而促进学术探讨?数据库内容可能引发哪些新的研究点?广泛运用于电商、检索平台的用户行为分析算法或可用于学术平台,通过分析平台访问日志来解决上述问题。

用户行为分析能使知识生产更有针对性、引导性。一方面,平台服务质量全面提升,用户的忠诚度能随之提高;另一方面,聚集兴趣点相近的人群,不仅能快速地挖掘史料,还可以减少资料搜集的个别性和零散性,从而对推进学术研究发挥实质性的作用。

三、着眼于史学应用化的历史共享平台

公共史学有三个维度:书写公众历史、让公众参与、写给公众看。^②公共史学诸维度的实现需要借信息网络发展的东风,催生历史共享平台建设。口述历史是公众加入平台建设的入口之一,是连接大众和学术的桥梁,是打破学术主体与公共主体之间二元对立格局的契机,是史学家履行社会责任的有效方式,更是技术进步为历史学发展提供的新路径。

(一) 公共史学与网络伦理

当记忆成为研究对象,口述历史成为普通民众表达历史观念的一种途径,即普通民众拥有所谓“共享的解释权”^③。这样让无声群体得以发声,使得社会充分了解该群体,从而谋求社会和谐。这也有利于综合各方面的意见,形成比较客观公正的历史认识,从而形成历史共识。正如从事残疾人口述项目的学者卡伦·赫什(Karen Hirsch)所说:“(口述历史)可以让某些群体发声,可以引领地方社会历史的新观点,还可以帮助对文化状况的深层次的理解并影响到每个人。”^④

口述历史在中国有良好的社会基础。除了有学界的参与,还有来自新闻出版、政协文史委等社会各界的强力推动,并被广泛应用于社会学、民俗学等多个学科领域。由于口述历史项目

^② 钱茂伟. 公众史学: 从根本上解决历史的应用问题[N]. 社会科学报, 2017-03-23(5).

^③ Frisch M. A Shared Authority: Essays on the Craft and Meaning of Oral and Public History[M]. Albany: State University of New York Press, 1990: 188.

^④ Hirsch K. Culture and Disability: The Role of Oral History[M]//Perks R, Thomson A. Oral History Reader. New York: Routledge, 1995: 214.

需要较大的人力成本投入,很多主持口述历史项目的机构都采取了征募、培训访谈员,分发口述历史项目等众多包的方式。崔永元口述历史研究中心、中华女子学院图书馆等机构都采用这种形式。他们的实践都为口述历史的公共化、网络化奠定了较好的基础。

随着信息技术的发展,网络社区已经自发地形成了公共史学的新场域。^⑤公共史学旨在思考历史学如何在当代社会中应用和发展,但目前历史学界与社会还处于疏离状态。学术界亟须正视在网络环境下历史学的传播问题,探寻一条适合历史学发展的新道路。学术界既要鼓励公众参与讨论和重新阐释历史,又要掌握公众历史观念生成中的话语权,以便引导历史研究健康、深入发展。

让公众参与历史研究共享平台建设是连接历史学界与社会的可行途径,但应注重平台参与者的协商和教育^⑥。

同时,历史研究共享平台需建构法律保障的预防性机制。随着信息网络的发展,线上公开征集资料、开放纠错和评论窗口成为可能。借助平台,民众参与到回顾历史的活动中,形塑共同的历史记忆。但是,口述历史可能涉及敏感、隐私等内容,网络传播更是增加了相关风险。因此,上述工作应是在有针对知识产权、人格权、传播权等法律保障的前提下进行。访谈前应该向受访者做好必要的说明,尊重受访人的尊严和自主性,通过充分协商,签署版权法律文件。拥有法律保障是获取受访者信任的前提,因而也是确保访谈内容接近事实,真切反映受访者内心世界的前提。但是保障范围的限制不宜过度扩张,不应影响学术研究的创新性。^⑦

历史共享平台建设应注重学术引导。口述历史的时间段限于一个人的生命周期,访谈成果会对现实产生影响,且这段历史尚无充分的时间沉淀,难以做出客观评价,因此更需要学术界的引领。历史共享平台建设应增强审核机制,提升知识来源的可靠性,鼓励利用平台从事学术研究,从而有利于校正公众历史讨论的盲目性。将文献史料、口述历史和学术研究三者结合起来,有利于形成人文研究的闭环,打造研究主题社区。历史研究或将成为提高民众文化素养、深化爱国主义教育的引擎。

(二) 以规范化为要点的口述历史线上垂直服务

根据口述历史自身的流程和要求、国家的相关规定和学术界的相关需求来策划口述历史线上垂直服务功能,促成历史研究共享平台建设。平台主体规划如下表。

表 1 口述历史线上垂直服务功能规划概览
Overview list of vertical services along the oral history

功 能	描 述
访谈培训	提供有关口述史的线上课程、培训视频、学者讲座、资质测评等。
访谈征集	搜集访谈资料的音视频广告,口述史料采集协议书、访谈员委托协议书网签等。
提供范本	口述者基本信息采集表,自动生成访谈基本问题、访谈介绍函、访谈成果清单(视频、音频、抄本、整理稿、照片、扫描件、访后感),自动匹配访谈前必读历史资料等。
编辑上传	录音的文字转换,扫描、照片、视频等的格式转换,音视频剪辑等。
审 核	方向:法律文件、导向审核、史实核对。内容:新颖、完整、深入。文字:错别字、可读性、篇章结构。

^⑤ 李红岩.如何应对“人人都是历史学家”的时代[J].探索与争鸣,2020(9):36-38.

^⑥ Neuenschwander J A. A Guide to Oral History and the Law[M]. New York and Oxford: Oxford University Press, 2014: 111.

^⑦ 杨祥银.美国口述历史伦理审查机制研究[J].史学理论研究,2016(2):86-97.

续表

功能	描述
纠错	用户可在线提交批评意见和修改建议,并提供证据和信息来源等。
评论	鼓励用户抒发阅读感受、分享类似经历等。
整理分析	编辑后台前台化,提供标准化分类、个性化标引、可视化呈现等。
辅助科研	关键词、多频词、标引项的统计分析;个性化投稿建议,链接期刊投稿平台;科研资助或项目申报渠道等。
融合出版	出版申请通道;人机合作编校;全自动排版;审稿流程一键查询功能,审稿进度更新提醒等。

线上垂直服务的要点是规范化,包括信息搜集、整理的规范化和审核上线的规范化。

学界对口述历史真实性的质疑,要求口述历史采集和使用的全流程更为规范。口述历史成果的归档可以根据我国档案局于2018年1月1日起实施的《口述史料采集与管理规范》(DA/T 59—2017)具体实施,并据此提供线上模板,包括口述史料采集协议书、口述者基本信息采集表、采集成果封面、著录标引字段、采集编号等。

口述访谈标题的拟定和摘要的撰写需要规范化,方便用户查找和阅读。标题分主标题和副标题两个部分,标题根据访谈内容拟定,主标题体现访谈的整体脉络或独特之处;副标题为“受访者姓名+口述的主要事件”。摘要参照贝丝·罗伯逊(Beth M. Robertson)《如何做好口述历史》一书中的建议,提供访谈内容的简单纲要,用以提示主题次序和录音各段落的相关信息,并陈述这段访谈的优点、焦点以及弱点(不要用带有强烈否定的词语)。^⑧

其他史料同样需要规范化的加工和标引。鉴于目前各历史数据库有著录标引规则不统一的情况,以及用户对元数据多元化的需求,平台需要制定一套基本的开放式的著录标引体系框架,形成积累语料库的底本。例如对保存版本、储存形式等内容加工的规定,以《中国分类主题词表》为基础构建标引体系等,都有利于规范语料库建设,增加自然语言识别的精准率,实现自动标引,从而减少人力成本和不规范问题。

口述历史平台建设对网络伦理和道德引领意义重大。鉴于信息传播渠道的拓宽和传播速度的加快,需要制定一套访谈员培训、优质资源推荐和资源上传的审核、发布的通用标准。这些规则的制定不仅需要学术界的探讨和参与,还需要政府有关部门强有力的支持。

(三) 平台共建与跨学科合作

历史研究共享平台的特点主要体现在科研辅助和便捷交互两方面的功能上,以研究主题为焦点,形成科学引领、规范化的网络社区。平台从功能到内容的实现,需要发动各学科研究者和各界民众共同参与,助推学术资料的积累和学术研究的进步。

资料的搜集与传播需要法律对著作权等各项权益的保障。中国传媒大学崔永元口述历史研究中心的口述史成果丰硕,却因为版权问题很难实现对公众开放,在传播和利用方面难免打了折扣。

资料整理和共享需要有图书馆和档案馆的加入。一方面,作为资料收藏的主要机构,图书馆和档案馆的资源互通与共享是打破“信息壁垒”的关键,其“开源”程度的高低决定了平台起步能否成功。2015年,国家图书馆的“中国记忆项目实验网站”向全国图书馆界发起共同开展记

^⑧ 贝丝·罗伯逊.如何做好口述历史[M].黄煜文,译.台北:五观艺术管理有限公司,2004:67-69.

忆资源抢救与建设的倡议,即是较好的尝试。另一方面,平台建设需要图书馆学、档案学界参与资料整理规则的制定。

口述历史已经在历史学、社会学、民俗学、新闻出版等领域有了跨学科的合作。随着数字技术的发展,各学科还需要与计算机技术结合。一方面以项目或问题为中心开展平台建设和学术研究,从挪用计算机技术向创新技术功能转变。^⑳^㉑ 另一方面需要计算机技术确保平台的网络安全和硬件安全。中华女子学院的“中国女性图书馆”收录了妇女口述历史成果,2013年建成中英文网站免费对外开放,有较可观的浏览量。但由于该数据库屡遭黑客攻击,于2018年被迫关闭,实属可惜。因此,口述历史平台建设还需要国家法律法规与计算机技术运用相配套,合法保护和有效利用个人信息^㉒,并确保算法的可控^㉓。

口述历史平台建设中音、视频剪辑的艺术性凸显。世界正处于媒体转型期,即从纸质媒体向数字媒体迁移,从以文字为中心向以形象为中心转换,大多数读者的阅读习惯和信息诉求已经改变。综述片、纪录片、影像志等多媒体成果,有利于历史知识的传播,形成品牌效应,激发公众的创作热情,推动平台发展。5G技术使得多媒体传播更为快速、高效。学术平台亟须正视从文献资料汇集向多媒体成果展示的转型。

通过跨学科合作,平台可以实现资源扩充、方法多元、关联扩展、形式多样,形成安全、实用、互动、规范、正气的网络社区。平台的建设和运营需要足够的持续的智力支持和资金支持,目前尚未形成有效的驱动链条。

四、结论:历史研究共享平台带来史学新发展

在数字人文飞速发展的新时代,将知识生产与知识保存打通,将史料搜集与历史研究打通,将学术研究与引领公众的作用打通,将文化遗产与公共文化打通等是发展趋势。历史研究数据库也应从史料型数据库向量化数据库,再向可持续扩展的历史研究共享平台发展。可持续扩展性通过共享、互通、交互来实现,主要体现为资源不断充实,知识关联不断丰富,知识服务功能不断优化。

习近平总书记在2019年全国“两会”上提出:“哲学社会科学工作者要多到实地调查研究,了解百姓生活状况、把握群众思想脉搏,着眼群众需要解疑释惑、阐明道理,把学问写进群众心坎里。”历史研究共享平台或可成为学者走出象牙塔,为群众做学问的一条可能路径。历史学者需积极响应党史学习的号召,通过平台搜集党史资料、分享党史学习和研究成果,或可形成解决跨界合作、融合发展、利益驱动问题的突破口,甚至可能形成新业态。

近代以来的资料浩繁,但是相对于其他历史时段的历史资料来说,全文数字化和自然语言处理较容易实现。从党史资料入手,能节省技术成本,循序渐进地实现功能开发,再逐渐辐射到社科其他领域,最终满足比较研究、量化研究、数字学术、公共史学、跨学科研究等领域对数字人文所提出的需求。如此,不仅丰富了资料获取方式,并可实现通过数据分析发现新的知识。

2021年5月29日启动的“当代文史数字书店暨‘四史’学习教育数字中心”项目是通过“四史”学习连接历史学与信息技术、连接历史学与公众学习的一次积极实践。以服务党史学习、党

^⑳ 台湾学术界近十年来提倡,人文学科与计算机技术的结合,不应局限于单向挪用计算机、统计学界既有的技术、算法,而是要开拓具有人文社科特色的新的分析工具和研究方法。项洁,翁稷安.数位人文在历史学研究的应用[M].台北:台大出版中心,2011.

^㉑ 邱伟云.我国台湾数字人文研究进程(2009—2017)[J].图书馆论坛,2020(7):9-19.

^㉒ 邵国松,黄琪.个人数据保护全球融合的趋势与挑战[J].上海交通大学学报(哲学社会科学版),2021(4):148-158.

^㉓ 崔聪聪,许智鑫.机器学习算法的法律规则[J].上海交通大学学报(哲学社会科学版),2020(2):35-47.

务工作为主的“学习强国”学习平台也积攒了较广的用户群体。未来或可以“学习”为契机,连接学术研究和公共文化领域,发动各界突破壁垒,集中解决资金和技术问题,激发知识经济和文化创意产业,使史学从基础学科向应用延伸,强化知识普及、文化引领等社会服务功能。

Construction of a Shared Platform for Historical Research: Taking Oral History as the Starting Point

LI Erling

(Institute of Contemporary China, Chinese Academy of Social Sciences, Beijing 100009, China)

Abstract: Oral history has been criticized for its veracity. Compared with other historical data, the collection, edition, and presentation of oral history needs to be more standardized to ensure its authenticity. Oral history, as an interdisciplinary medium that connects academia and the public, has the advantage of becoming the point cut for building a shared platform for historical study. The platform can meet the needs of “digital humanities” in comparative research, quantitative research, digital scholarship, public history, interdisciplinary research, and other fields. It will enrich the way of obtaining data and can discover new knowledge through data analysis. Building a public cultural exchange platform with sustainable expansion of resources, knowledge connection and knowledge service can extend history from basic disciplines to applications, strengthen social service functions as knowledge popularization and cultural guidance.

Keywords: oral history; stereoscopic data database; shared platform for historical study; social service; new format

(责任编辑:何菲)