

专业史学工作者的一般要求。换句话说,没有基本技能,缺乏“搜商”,新样态的数字化、信息化史料,也难以和史学研究效率的提高结合在一起。伴随大数据时代的到来,历史资料的电子文本随之增多是不争的事实,历史学工作者查阅、搜集、整理、占有史料过程中,新样态史料所占比重会越来越大,通过“搜商”推进阅读速度,提高工作效率,进而提高学术研究效率是非常现实的问题。这在实践中已经是一个完整的逻辑链条,关键环节是“搜商”。

当然,希望以数字化、电子化形态的史料完全取代传统的纸质文本史料,期待“超纸质文本”时代的到来,尚需假以时日。关于数字化史料及其阅读,许多人提出异议,认为纸质文本无可替代,认为电子阅读缺少了“读书”的意味。其实,笔者以为,今天所见某些新样态史料,许多内容仅仅“数字化了”的“故纸堆”,原始意义上的“故纸堆”只不过在数字化时代改变了存在形态,变化了外在表现形式,并未改变内涵价值,“故纸”的性质并未因“被数字化”发生改变。面对日新月异的数字化技术,面对层出不穷的数字化产品,历史学工作者改变

(作者系哈尔滨师范大学副教授,历史学博士)

阅读习惯是必要的。虽然不可以设想让拥有深厚钻研故纸堆传统的史学工作者变成“网虫”,但如果依然固守原有的研究模式、查阅资料的模式,也存在落伍时代之嫌。

本文所言数字化史料与史学研究效率的提高,仅仅以技术生成为视角,并非否定史学悠久的传统。比如,数字化技术为搜集整理史料创造的便利条件有目共睹,但真正提高史学研究效率依然不可漠视对史料的辨伪、考证、甄别、比较等历史学不可或缺的研究方法,不仅没有被消解,反而在新的社会条件下,彰显出比之先前更重要的价值。离开了这些研究方法,史料的可信度必然随之降低,研究效率无从谈起。因此,我们必须强调,数字化史料、新样态史料为提升史学研究效率创造了必要条件,但并不是充分条件。我们强调的是,历史学工作者在掌握了历史学研究基本方法后,尽可能多地利用数字化史料,进而提升研究的效率。

[参 考 文 献]

[1] 张晓校. 新样态史料与历史书写[J]. 北方论丛 2015(4).

[责任编辑 张晓校]

大数据时代对历史研究影响刍议

李红梅

(哈尔滨师范大学 历史文化学院 黑龙江 哈尔滨 150025)

[中图分类号]K0

[文献标识码]A

[文章编号]1000-3541(2016)02-0077-03

DOI:10.13761/j.cnki.bflc.2016.02.015

历史学的进步与科技发展紧密相关。从这个意义上说,历史研究当前正面临一个新的发展契机,这就是大数据时代的来临。大数据已经成为世界的本质和时代特征,对历史研究产生深刻影响也是不争的事实。

首先,从历史证据角度观察,大数据带来史料挖掘和阅读方式的革命,这是大数据对历史学发展最大、最重要的影响。历史学是基于证据的学问。如何获得历史证据——史料,是历史研究中最重要,也是最基础的部分。按照现代泛史料的

[收稿日期]2015-12-19

共识,人类生活各个时期各个层面留下的所有痕迹,都是复原历史的证据。史学界基本能达成一致的分类标准,大致将史料分为文字史料、实物史料、口碑史料和声像史料等。因为史料种类繁多,保存分散,历史研究中搜集史料,成为最耗时、耗力的工作,费尽心力可能只得到只言片语。历史研究要求对相关史料“竭泽而渔”,但在传统纸质文本时代,史学家很难将史料“一网打尽”亦属“常态”。网络时代、大数据时代,数字化技术的进步,改变了史料的原始存在形态,传统存在方式的各种史料,借助信息化、数字化技术生成了新的存在样态——电子文本或数字化文本。新样态史料出现于20世纪90年代,最先进入人们视野的是文本史料数字化。新样态史料革命性地改变了挖掘和阅读史料的形式,其为历史学工作者尽可能快捷地穷尽史料创造了条件。

第一,直接查阅相关数据库寻找历史证据。目前,历史研究中最重要数据库是文本文档生成的PDF格式的数据库,这些数据库既有第一手史料,也有各式研究成果和已有出版物等。电子文件柜式的数据库形式多样,内容既有综合性的,也有专业性的。更先进的数据库是存储海量数据的大型数据化系统。利用数据库进行历史研究的优点非常突出:相关内容集中,寻找史料方便快捷,事半功倍;数据库的内容不需要进行史料的外考证,尤其是PDF格式的数据库,这个优点尤为明显;可以比较快捷地进行学术史梳理。

第二,充分利用搜索引擎在网上寻找史料。除了众所周知的、较重要的数据库外,在比较专业的研究中,还有很多相对比较小众的数据库,或者比较分散的史料,需要通过全网搜索的方式挖掘,可达到穷尽资料之目的。尤其现在海量存在的大数据,给研究工作提供了更多可供寻找的资源。这种挖掘史料的方式比使用数据库难度大,考验史料搜集者的数字化技术的基本技能的把握程度。如关键词的筛选、搜索路径的确定等。和传统史料搜寻的方式方法相比,网络时代、大数据时代的新样态史料搜集主要在虚拟空间内进行,不依赖数字化技术达到目的。

第三,积极利用各种数字化技术生成的声像史料,进行现代意义的历史研究。传统历史研究运用的史料,不论文字,还是实物、图片,都是静态的,终极研究成果以平面文档形式呈现。网络阅读、电子阅读,可以包括利用动态的声像资料在内的新样态史料,这是现代历史研究激动人心的变化,历史可在某种程度上活生生地展现在世人面

前。当然,这种全新的史料如何运用,成果如何体现是一个新问题,需要尽快确立引征规范。时下,真正运用声像史料作为最重要史料研究历史还是一个有待开垦的新领域。

第四,借助电子技术手段生成新史料。生成史料在技术上有简繁之分。简单的如在相关问题的研究中,借助现代的电子技术及其相关软件,将文本资料中文字和数字等记载,转换成音像等新形式史料,生动形象说明问题。复杂的做法,是历史学家利用计量和统计的方法,按照研究问题的需求和已有数据之间的关联,由原始数据计算生成新的数据,即生成数据。此前计量史学已经进行过这种探索。大数据时代,数据挖掘具有更大的必要性。史学家既要大数据的本来意义中探寻历史真相,又要挖掘数据的潜在意义,从中发现新知识,此为大数据的更大价值。

其次,从历史叙述的角度考察,大数据有助于复原真实的历史。大数据突破了地域性和国家性的限制,很大程度上实现了资源共享。傅斯年先生曾有“上穷碧落下黄泉,动手动脚找东西”之名言。今天,历史学家有了在网上“动手动脚”的广阔空间,许多问题的研究可以借助相关史料,多方求证,互相对印证,比之传统史料更加得心应手,无疑对历史研究的影响是积极的。

第一,有利于搭建交流的平台。史料过度难寻和不充分的时代,即使史学家关注共同的题域,也会因搜寻史料的片面性,造成历史叙事在某种程度上是自说自话,互不相搭界。大数据时代共享史料的范围逐渐拓宽,历史研究的前提和基础有了相似性,研究成果可以找到交流和承接的平台。当然,历史研究有不同的理论和解释前提、不同的叙事框架,彼此间的话语系统永远不会完全相同,但因为史料限制的突破,成果交流和思想碰撞的机会增加,故步自封和自成一体的系统容易打破,充分的交流自然会促使历史研究进步。

第二,有利于叙述更真实的历史。复原真实历史的前提,是寻求到所有可能存在的资料,形成完整的证据链。越来越多和越来越精确史料的获得,使得历史叙述不再停留于含糊、推测、大概的程度,历史的面貌逐渐清晰。王国维先生提出的两重证据法,大数据时代真正有了实践的可能。历史研究的课题宏观,也好微观也罢,因为充分的史料,其复原的历史真实性更大。当然,从认识发生学的角度看,不可能得到完全真实的历史,毕竟历史现象复杂,复原全部真实的历史还是难以企及的梦想,但大数据史料使史学家接近历史真相

的步伐加快。

第三,有利于历史叙事的多样性。传统史学叙事,更多的是关注社会上层和政治领域;新史学将叙事主题、叙事对象扩展到经济、社会、文化等领域,既可以是宏大叙事,也可以关注普通人的历史。通过大数据技术及其大数据搭建的平台,史料不仅充盈,种类繁多(包括音像资料),而且获得路径多种多样。大数据时代,历史叙事领域会极大拓展,研究领域可以囊括宏观到微观的多个层面。大数据使个人在真实世界的行为得到准确精细完整的记录,人类生活的各个侧面都可以进入研究视野。在海量的数据中,限定范围、象限,可以形成一个个小拼图,无数小拼图,就可以形成大拼图。史学家有更强的能力去发现多侧面和多层次的历史。

最后,从历史解释角度观察大数据的影响。历史解释是寻求历史现象背后的因果链条,寻求历史发展的脉络、寻求历史对现在和未来的启示与意义。寻找历史现象背后的本质,是人类理性发展的要求。复原的历史真实清晰,证据链充分,史学家解释历史时就会有更辽阔的视野,结论会更令人信服有说服力,这是毋庸置疑的。

历史研究最大的意义是人类对自身行为的探索和理解,从这个意义上说,历史研究中不同的观点和解释可能是史学家最大的贡献。历史常看常新,每一代人重新书写历史的意义就在于此。因为历史解释具有主观性,否认大数据时代历史研究已经和可能发生的巨大变革,显然是不正确的。毕竟面对海量数据的庞大和公开的资料群,史学家会不断提高对待历史证据的公正程度。为了主观利益和目的研究历史,随意“打扮”历史,漠视证据,不公正地解释证据,会受到越来越多的关注和谴责。犹如在史学研究中,参与者都有了知情权,过于偏狭的主观性、与公认的道德和价值观完全背道而驰的主观性,会因为更多的监督而有所收敛。

最后,在讨论大数据对历史研究的影响时,应对大数据之于历史研究的影响略陈己见:

第一,大数据对历史研究最大的影响是史料新样态化,最终的结果是史料群空前庞大,历史学研究史料总量及其可选择性日益增多,彰显了大数据时代信息数据的海量特征,为历史研究提供

了前所未有的便利条件。但历史研究工作本身的性质、追求并没有改变,历史研究最基本研究方法也没有变化,只是面对新样态史料,史学研究者除了传统的研究技能、基本技术之外,数字化、信息化基本素养是必要的,唯有如此,才能使历史学研究在大数据时代表现出时代特征,才能使大数据在史学研究中发挥最大的效用。

第二,数据化与历史研究有机结合尚存在较大的上浮空间。因数字化技术生成的新样态史料已经被越来越多的史学工作者所重视、所利用,一些专业化、专门化数据库的建设也取得了巨大成效。大数据时代,数据库将在历史学研究中发挥更大的作用,但时下的数据库建设还存在诸多遗憾,远远无法满足历史学各个学科领域研究的需要。作为历史学研究的基础资料——史料并不稀缺,稀缺的是如何尽快使这些林林总总的史料转化为新样态史料,这也是历史学对大数据时代的呼唤。与历史学相关的各种数据库的建设,既需要仰赖信息专家的技术贡献,历史学家的智慧也不可或缺。如何实现这种跨界合作,是广大史学工作者理应思考的课题。在此,我们不得不重申历史学工作者的“大数据意识”,即在大数据时代,历史学工作者在恪守传统的同时,不忘记数据库为代表的“史料宝库”,通过数据库获得更多、更精准的史料。

第三,历史学的数据化与数据化了的历史学。无论对历史学有何种解读,历史本身承载着各种信息,正因如此,史料方能与数字技术结缘,历史所承载的信息也具备了“数据化”的前提。诚然,将史料理解为数据似乎在彻底扭转史学观念,各种史料将被“被数据化”一定是大势所趋。历史学工作者在获取各种史料时,为“数据”四处奔波,这种从在“故纸堆”中寻觅、爬梳,到在大数据中搜索的转型,不仅是技术手段的更新或转换,而且会因史料获得的便捷,产生新的飞跃。可以预见,一个“数据—史料”的时代,正在向我们走来。

大数据改变着人类的社会生活,必然改变历史生存与研究的现状。令人欣慰的是,这种改变产生的各种效应是积极的。对大数据充满期待,似乎应成为历史学工作者的一种期待。

(作者系哈尔滨师范大学副教授)

[责任编辑 张晓校]