

助力历史学繁荣发展

## 把握数字技术对历史研究的赋能作用

王涛

习近平总书记指出：“重视历史、研究历史、借鉴历史是中华民族5000多年文明史的一个优良传统。”作为一门古老的学问，历史学生命力的延续源于自身的开放性。20世纪60年代，有历史学家呼吁历史学要与地理、生物等学科结成新同盟，展开跨学科研究，为历史学带来多样性的发展。当今世界，科技进步日新月异，互联网、云计算、大数据等数字技术深刻改变着人类的思维、生产、生活、学习方式。在数字技术不断迭代的背景下，今天我国历史研究工作者也在积极探索与数字技术融合的路径，努力顺势而为、力争乘势而上，让历史研究更加多元化，进而推动历史学繁荣发展。

数字技术赋能历史学，涵盖史料挖掘整理、历史问题研究、研究成果发布传播等各个环节。数字技术可以帮助历史研究工作者收集、量化、解释和分享历史资料，揭示其中蕴含的信息。同时，当数字技术成为自身技能的组成部分，历史研究工作者也会自觉从算法等技术角度看待历史问题研究，如此一来，历史研究的路径就会得到拓展，历史书写的方式也越来越多

样。此外，数字技术还能用可视化等方式将历史研究成果呈现出来。从这个意义上说，数字技术能给历史学带来新的研究视野和方法，让历史研究工作者具备了更多的思维模式。

历史史料使用的广度与深度得到拓展。历史研究需要立足史料展开。在数字技术的赋能下，史料的使用效率得到了提升。数字技术帮助历史研究工作者深化了对史料的认识，史料不再局限于文本材料，而是出现多元化的趋势，图像、物质材料等在完成有效的数字化之后，都能够进入历史研究的视野。历史研究工作者可以足不出户使用各种类型数据库，还能利用检索技术快速定位获取目标文献。同时，一些此前无法获取的史料，比如由于年代久远或者因为外在因素（自然灾害、人为破坏等）遭到损坏的史料，还有容易损坏的脆弱历史文献，都可以用机器学习的方式得到重建。再有研究团队利用透视扫描技术，结合机器学习算法，快速有效提取获取书信等文献内容，使其成为历史研究的材料。人工智能还能帮助历史研究工作者翻译古代文字，包括

甲骨文、西夏文、残缺希腊语铭文等在内的古文字都在数字技术的协助下得到了一定程度的识别与释读，扩充了古代文献的范畴。此外，之前由于数量太大无法得到有效使用的材料，也能在数字技术的协助下通过算法提炼出具有学术价值的信息。比如，欧洲中世纪的贵族家族留下的大量纹章素材，是研究近代早期欧洲文化与社会状况的重要资料。但由于纹章材料过于庞杂，其形制与使用场景也不尽相同，研究难度极大。数字纹章项目使用机器学习以及语义网技术，实现了对纹章的自动描述与识别，使相关研究得以顺利开展。

历史研究方式得到改进。在得到数字技术赋能后，历史学的开放性让历史研究工作者可以充分利用不同学科的理念和方法来推进历史研究，比如可以利用进行空间分析的地理信息系统、社会学中的社会网络分析方法、计算语言学的主模型模块等进行分析研究。这些跨学科的方法与历史研究结合能够发挥学术效能，甚至能获得具有颠覆性的发现。比如，国内有学者利用数字技术建立了“数字概念史”的方法论，基于人机共读，梳理了中

国近代海量文献中思想状况的演变情况，追踪“道”概念的变迁过程，得出了超越传统认知的新发现。同时，在运用数字技术的情况下，历史研究的工作流程有了更多合作需求。数字技术日新月异的发展意味着越来越多的数字技能需要被掌握，但历史研究工作者不可能掌握所有数字技能，不同领域的学者进行合作研究就成为必然选择。这不仅让数字技术赋能下的历史研究课题在形式上具有团体协作的必要，而且这种跨界融合的研究方式也在实质上拓展和深化了历史研究。当下，大语言模型如何接入历史研究之中是热门议题。单纯从人工智能的角度看，在方法论上使用机器学习的方法，已经说明人工智能可以与历史研究紧密结合。国内已经有高校开发了基于中国古代文化典籍的大语言模型，能够实现点校、实体提取、翻译等功能，学术意义重大。当前的问题在于，历史研究工作者如何评估人工智能大模型生成的文字内容，如何让人工智能参与的历史书写具有更好的学术性并更好地遵循学术规范，如何与时俱进建立有效的同行评议机制。这些问题并不会阻碍历史学家使用

人工智能，但需要我们让人工智能更加规范地扮演学术助理的角色，推动学术研究中人机协作稳健发展。

历史研究成果的发布传播形式得到丰富。历史研究工作从课题选择开始，经历漫长而艰辛的研究过程之后需要将成果发布出来，整个研究工作的流程才能形成完整的闭环。传统的学术发表体系具有相对封闭的特性，从形式上看，论文、专著、报告等形式比较单一，能够呈现的内容也十分有限，制约了研究成果社会效益的发挥；从效率上看，单一的出版媒介不利于历史研究成果及时被学术圈外的大众接受。数字技术让学术成果的发布传播摆脱传统方式的束缚，让历史研究在内容上更加丰富、在受众上更加多元。比如，敦煌研究院推出的“数字敦煌”项目，充分挖掘数字技术对知识传播的作用，用数字技术对敦煌石窟、壁画等文化遗产进行数字化复原，既在形式上更加多样，也让敦煌文化研究成果的传播更加有效。可以预见，数字技术的不断进步，必将进一步拓展历史研究成果的发布传播渠道。

（作者为南京大学历史学院教授）

提供新机遇、新视角、新工具

## 发挥好数字历史研究方法的重要作用

王旭东

学术的发展离不开创新。习近平总书记在强调“使中国特色哲学社会科学真正屹立于世界学术之林”时指出，要“不断推进知识创新、理论创新、方法创新”；在对我国广大历史研究工作者提出殷切期望时强调，要“着力提高研究水平和创新能力”。当前，信息革命的时代潮流浩浩荡荡，有力推动人类社会向数字化、网络化、智能化方向迈进。以大数据、云计算、人工智能等为代表的数字技术不仅深刻影响着人们的生产生活，也为哲学社会科学研究带来了机遇和挑战。近年来，数字史学、信息史学等蓬勃兴起，充分表明数字技术为历史研究提供了新机遇、新视角、新工具。其中，从方法论层面出现的数字历史研究方法，即用数字技术收集、处理、分析和解释历史数据，对广大历史研究工作者收集、提取、研究各类历史资料具有变革性意义。

数字历史研究方法，是一种将数字技术与历史研究相结合的新兴方法。它以数据驱动为基础和前提，利用计算机技术和数字工具来收集、处理、分析和解释历史资料，对历史事件提供趋势性的定量或定性分析，使历史研究更加准确、全面和深入，能够助力历史研究工作者开辟新领域、提出新问题，进而对历史事件和历史过程的复杂性提出新见解。具体而言，数字历史研究方法包括以下几个方面：一是数字化文献建立，即将历史文献和档案数字化，以便于存储、检索和分析；二是文本挖掘数字化，即利用自然语言处理和机器学习技术，对历史文献进行文本分析和语义解析，从中提取有关历史事件、人物和主题的信息；三是研究分析数字化，即通过构建历史人物、组织、事件等的网络模型，分析它们之间的关系和影响，揭示历史事件的复杂性和动态性；四

是数据可视化，即利用图表、地图、知识图谱等可视化方法，将历史数据以直观形式进行展示，帮助研究者理解和解释历史事件；等等。

当前，数字历史研究方法正在迅速推广和应用。许多国家和地区的学术机构和研究组织制定了数字历史计划，促进该领域的合作和交流。不少图书馆、档案馆和研究机构都建立了数字化历史资源库，并提供在线访问和使用的平台，为研究人员提供大量的历史档案和文献资源。同时，一些数字化历史工具和软件也得到了广泛应用，如嵌入式OCR技术（光学字符识别）、自然语言处理技术、数据可视化软件、历史地理信息系统平台等，这些技术和工具助力历史研究更加高效、精确和可视化。

面向未来，数字历史研究方法必将随着自身的发展以及与其他学科的不断融合而发挥更大作用，进而为

推动新时代历史学繁荣发展作出贡献。例如，人工智能的集成应用将发挥更大作用。机器学习和自然语言处理技术将提供更加智能的数据处理和分析能力，不仅可以帮助广大历史研究工作者自动完成文档分类、转录等任务，还可以从大型数据集里生成假设，从而在海量历史数据中探索出新的分析模式和研究路径，由此获得对历史事件和历史过程的新见解。又如，数据的多样性、开放性和可获取性将创造更多研究机会。未来，越来越多的历史研究资料将以开放数据的形式提供，这将促进学术合作和跨学科研究，实现数据、方法等的共享，为个体研究者跨时空参与相关研究提供可能。此外，数字历史教育也将进一步发展，从而吸引更多人通过在线课程和教学工具参与各类历史研究。再如，创建交互式、可视化情景能力的提升将开辟更多研究平台。随着增强现实、虚拟现实等技

术的升级和普及，重建历史环境和景观将变得相对容易，广大历史研究工作者和公众能够以前所未有的沉浸式方式与过去互动，增强历史研究的沉浸式体验。

在信息时代的数字化转型中，数字历史研究方法正在用新兴技术改造古老的历史学。需要明确的是，无论数字技术怎么发展，我们都要始终保持对历史资料的尊重，避免过度数字化而忽视历史事件的复杂性和多样性。更重要的是，要时刻清醒地认识到，无论数字历史研究方法发展到何种程度，它始终是一种技术工具性质的方法。在使用数字历史研究方法时，作为研究主体的历史研究工作者要坚持唯物史观。唯有如此，才能科学有效地利用持续更新的先进技术手段，助力新时代中国历史学繁荣发展。

（作者为中国社会科学院世界历史研究所研究员）

坚持运用辩证唯物主义和历史唯物主义

## 正确认识和科学对待史料数字化

董灏智

史料是历史研究的基础和前提。尽可能系统地搜集、整理、分析有关史料是历史研究的基础性工作。习近平总书记指出：“要坚持用唯物史观来认识和记述历史，把历史结论建立在翔实准确的史料支撑和深入细致的研究分析的基础之上。”关于史料的内涵，学术界比较认可梁启超给出的定义：“史料者何？过去人类思想行事所留之痕迹，有证据传留至今日者也。”时至今日，史料的分类越来越细，除了文献史料、实物史料、口述史料等常规史料，还包括影像史料、音频史料等，特别是数字技术的快速发展，不仅极大推动史料数字化构建，也在生成大量原生的或外延的数字化史料。比如，史料数据库的建设为历史研究工作者检索、复制文献史料提供了极大便利，人工智能技术等对古籍文字进行识别排序进而激活古籍生命力，等等。同时要看到，随着史料数字化发展向纵深推进，在历史研究中也产生了“技术万能主义”等错误认识。因此，无论是面对史料数字化还是数字化史料，我们都要坚持辩证唯物主义和历史唯物主义，真正让数字技术推动历史学向更为宽广的领域

发展。

在历史学发展史上，始于20世纪初期的西方“史料革命”倡导扩大历史研究范围，由政治史扩展至社会史、经济史、文化史、生态史等。在这一过程中，史料的范围亦随之扩大，不再局限于官方档案和传世历史文献，还包括书信、日记、碑刻、家谱、契约文书、账本、回忆录等。尤为重要的是，20世纪50年代计算机技术被引入历史研究之中，使用计算机技术处理历史数据，把各种形式的文字史料转化为数字信息并进行加工、处理、存储，开启了史料数字化的进程。史料数字化早期主要采用直接扫描的方式，将文献史料逐页传输至计算机内，存储为图像文件，不仅保存了史料的原貌，还使零散的史料集中起来。在此基础上，计算机技术将文献史料转换成文本形式，建立了可以全文检索的数据系统，史料数据库由此出现，从技术层面进一步推动了“史料革命”。同时，互联网技术的发展加快了数字化史料的传输速度、拓宽了史料的获取渠道，促进各种“史料数据库”的开发。就我国而言，史料数字化开启于20世纪90年代，其发展脉络与西

方较为相似，也是采用先扫描、后文本化的形式将“二十四史”“十三经注疏”“诸子集成”“四库全书”以及近代报纸、杂志等史料大规模地数字化、信息化，并利用网络技术开发了各种类型的史料数据库。在近30年的时间里，中国国家图书馆、高校和科研机构等先后建成了数以百计的全文检索、高级检索等多条件检索的数据库。以“中国基本古籍库”和“全国报刊索引”为例，前者收录了自先秦至民国时期的历代文献1万多种，共计17万卷，全文约17亿字，数据量约330G，是中国最大的历代典籍总汇；后者收录的报刊数据可追溯至1833年，至今已累计收录报刊5万余种、文数5000多万篇，年更新数据量超过500万条。此外，大量的国外史学数据库引入国内，不仅促进了中国的世界史研究，更与国内史料数据库建设共同推动了史料数字化进程。

史料数字化使史料呈现出几何级数增长的趋势，极大扩充了历史研究工作者可以使用的史料规模，在多个方面对历史研究产生了深远影响。第一，无论是史料数字化，还是数字化史料，都能让历史研究

工作者占有越来越广泛和丰富的史料，有助于实现研究对象从“宏观”向“微观”扩展。历史研究工作者利用数据库中的海量文献，既可以对政治史、军事史、思想史等进行宏观研究，也可以对经济史、社会史、文化史等进行微观研究，通过研究日记、书信、账本等私人文献，折射出大历史背景下个人的生活与思考，进而理解历史事件的复杂性，在一定程度上弥补了宏观研究的不足。第二，史料数据库改变传统的历史研究方式。史料数据库将众多史料融为一体，其中的全文检索、高级检索，为历史研究工作者快捷查找史料提供了便利。数据库使史料搜集方式从抄史料、做笔记等转变为便捷化的智能检索，在获取史料的途径、数量以及效率方面均得到了空前提高。第三，史料数字化进一步推动跨学科研究。比如，对历史研究工作者而言，不但是要掌握基本的数据挖掘、数据分析等方法，还要了解数字技术的最新动态，从而实现历史研究与数字技术紧密结合。此外，针对数字化史料的研究，还需要借助人类学、民族学、民俗学、人口学、统计学、生态学等学科的研究方法，通

过对史料的定量、精准分析，得出相对客观的结论，能够促进历史研究视角的多元化。

也要看到，史料数据库在给历史研究带来极大便利的同时，其局限性也日益凸显。史料数据库只是历史研究的辅助工具，并不是历史研究的全部，检索也无法代替阅读和思考。若无扎实的史学基础训练和对整体历史脉络的把握，经由史料数据库检索得出的大多是完整的信息，而仅仅依靠这些不完整的信息进行研究就会影响历史研究的整体性、系统性。同时，虽然史料数据库使用便捷，但如果历史研究工作者没有深入考察史料来源和真实性，便不能完全信任数据库中的史料。特别是文书、账本、日记、书信等私人史料或民间史料，若无相关史料的佐证，是不能视为权威的一手资料的。面向未来，广大历史研究工作者要加强史学训练，切实提高史学基础和问题意识，在研究中始终坚持辩证唯物主义和历史唯物主义，秉持正确的史料观念，科学合理利用包括史料数据库、人工智能大模型技术等在内的各种数字技术。

（作者为东北师范大学历史文化学院教授）

## 数字化赋能新时代历史学发展

习近平总书记指出：“当今世界，科技进步日新月异，互联网、云计算、大数据等现代信息技术深刻改变着人类的思维、生产、生活、学习方式”。在新兴数字技术的赋能下，古老的历史学焕发出新的生机。如何把握数字技术对历史研究的赋能作用？如何发挥好数字历史研究方法的重要作用？如何正确认识和科学对待史料数字化？本期学术版围绕这些问题进行探讨。

——编者