

数智时代与历史研究

【编者按】习近平总书记高度重视历史学习和历史研究,希望广大历史研究工作者着力提高研究水平和创新能力,推动相关历史学科融合发展。同时,习近平总书记高度重视人工智能在国家未来发展中的地位 and 作用,多次指出其为赢得全球科技竞争主动权的重要战略抓手。当前,以 ChatGPT 为代表的人工智能数字技术蓬勃发展,必将深度浸入人类社会生活与学术研究的各个领域。在可见的将来,人工智能与历史学的结合既是实现历史学创新的生长点,也是中国历史学加快构建中国特色历史学“三大体系”,立世界史学之潮头的重要支撑。为回应当前人工智能技术进展对历史学的冲击和影响,深化这一形势下的数字史学研究,本刊特邀五位相关领域专家,就生成式人工智能之于历史研究的机遇与挑战、数字史学未来走向、历史知识生产模式的变革及史学理论与人工智能的双向互动等问题展开讨论,以引起学界关注数智时代的历史研究的变化与动向。

生成式人工智能之于历史研究的机遇与挑战

王 涛 (南京大学历史学院教授)

面对日新月异、高速发展的大语言模型,生成式人工智能愈加频繁地出现在人文学者的视野之中。历史学家的态度出现了两个极端,可谓泾渭分明。一种态度对生成式人工智能不屑一顾,揪住生成内容出现的一点小错误就大做文章,认为大语言模型完成的历史书写不名一文。另一种态度则积极拥抱大语言模型,鼓吹人工智能的领先优势,认为消极应对是徒劳的,因为学术研究必将进入一个由生成式人工智能重新定义的时代,^①历史学家也应该“为人工智能预留发展空间”。^②按照这些乐观主义者的说法,不使用大语言模型的学者将会错失良机。^③

这是两种带有情绪的态度,要么贬低生成式人工智能之于历史研究的可取之处,要么无视其中存在的局限性。在早前的文章中,笔者认为面对大语言模型的正确态度,应该是将它作为开展历史研究工作的个人助理:合理利用它在文本生成方面的优势,提高研究者的工作效率,但也要对其结果保持批判吸收的能力。^④近几个月,大语言模型在数量级与算力上又取得新进展,需要我们持续关注和讨论相关话题。

① Fiona Fui-Hoon Nah, Ruilin Zheng, Jingyuan Cai, Keng Siau, and Langtao Chen, “Generative AI and ChatGPT: Applications, Challenges, and AI-human Collaboration”, *Journal of Information Technology Case and Application Research*, Vol. 25, No. 3, 2023, pp. 277 – 304.

② Darrell Meadows and Joshua Sternfeld, “Artificial Intelligence and the Practice of History”, *The American Historical Review*, Vol. 128, No. 3, 2023, p. 1348.

③ Matthew Hutson, “Could AI help you to write your next Paper?”, *Nature*, Vol. 611, No. 7934, 2022, p. 192.

④ 参见王涛:《大语言模型时代历史书写的路径与局限》,《澳门理工学报》2023年第4期。

一、人工智能赋能历史研究

历史学家的工作内容说起来既简单又复杂。简单在于,历史学家的工作内容总结起来就是读书、思考、写作。复杂在于,每个环节都可以持续展开,细分出不同的方面。比如,读书的环节其实是研究的准备工作,涉及读什么、去哪里读、如何读等更多细节问题。在试图结合数字史学方法的研究工作中,则会加入建制数据库等更复杂的工作流程。^① 史学实践的复杂性,为人工智能的参与预留了空间和想象力。

历史学家使用人工智能协助标注史料、查询笔记、总结文献主题、史料分类等工作,都是在助力而不是阻碍学术研究。在具体研究实践展开的环节,人工智能同样能够发挥效能,帮助历史学家更充分利用史料,提高研读文献的效率。特别是面向未来的历史研究,史料更多以“天然数字化”(born-digital)的面目呈现在历史学家面前。这也意味着此种类型的史料数量巨大,在处理上将消耗大量时间成本。比如,电子邮件是重要的历史资源,但如何对这些数据进行结构化建制,对史学家提出了极高的要求。要想让电子邮件得到有效利用,充分挖掘其学术价值,不仅需要史学家对电子邮件的内容有大致了解,还需要正确提取电子邮件的各种关联信息,比如涉及的时间、地点、人物、事件等。英国布里斯托大学(University of Bristol)的研究团队,利用人工智能技术,开发了电子邮件语境发现工具(email contextualisation discovery tool),让历史学家能够便捷地对邮件内容进行提取,方便后续的研究。^② 山东大学数字人文实验室在 2024 年初发布文翰边疆古籍大模型,能够实现文本翻译、实体识别、文本摘要、文本标注标点及词法分析等自然语言处理任务,为相关领域的专家提供了更强大的学术支撑。^③

随着人工智能技术进一步升级,更多机构加入研发,可供选择的大语言模型不断涌现,在功能的完善度及结果的准确度上,生成式人工智能的表现越来越优秀,在历史书写方面对历史学家的冲击似乎越来越大。最新升级的 GPT-4,以及谷歌公司推出的竞品 Gemini,最大的改进是将人工智能推进到能够处理多模态对象的高度,从而能够满足历史学家应对多样性史料的需求。在应用层面,大语言模型发展出的另外一个趋势就是用户端的友好化,即功能调用的极简:④在使用流程越来越简单的同时,实现的功能不仅没有打折,反而会带来使用效率的提升,从而增加人工智能技术在历史研究领域被成功利用的可能性。目前的最新进展是,生成式人工智能不仅能够辨识图片,对原始文献进行精度更高的 OCR 处理,还允许用户上传自己的文献,将大语言模型打造成专属的“领域专家”。

在开放 GPTs 权限后,用户可以定制个性化的“功能模块”。笔者曾经在 GPT 后台建立了一个 PDF 分析器的应用程序。令人惊艳的是,制作应用程序的过程,不需要任何编程语言的知识,直接用自然语言提出需求即可,大语言模型不仅能够给出正确的反馈,还会非常贴心地推荐程序的 Logo 设

① 梁晨、李中清:《从求实到求是:数字史学的价值与追求——兼与成一农教授讨论》,《清华大学学报》2022 年第 3 期。

② Stephanie Decker et al., “Finding Light in Dark Archives: Using AI to Connect Context and Content in Email”, *AI & SOCIETY*, Vol. 37, 2022, pp. 859 – 872.

③ <https://mp.weixin.qq.com/s/GJb3uj7TsjvgZmOVzNvXzQ> [2024 - 03 - 25]

④ Sam Sedaghat, “Success Through Simplicity: What Other Artificial Intelligence Applications in Medicine Should Learn from History and ChatGPT”, *Annals of Biomedical Engineering*, Vol. 51, 2023, pp. 2657 – 2658.

计(解释设计理念)、功能设置(实现哪些功能)和用户定位(面向大众还是专业人员,决定交流的语言是通俗易懂还是学术高雅)。更重要的是,在应用程序的设计环节,用户可以上传指定的文献,让应用程序接受训练,也就是说可以预制应用程序的出厂设置,让其一开始就以某一特定领域的专家起步。比如,如果我们给应用程序上传德国第二次世界大战(以下简称“二战”)史研究的文献,它将成为擅长分析二战历史的学术顾问。

人工智能在“识图”能力上的提升,对于拓展历史研究的素材,显然具有积极推动作用。笔者曾经随手上传了一张讽刺漫画,没有给出任何提示,只是询问 GPT-4,希望它能够提供一些背景信息。在极短的时间内,它就给出令人满意的答案。大语言模型获得多模态的处理功能,会给予历史研究巨大推动作用。随着史学理论不断发展,越来越多的历史学家意识到,史料的范畴将会突破单纯的文字材料,向更多元化的维度拓展,包括图像、声音、影像、人造物等,也即所谓“非文本资料”(non-textual source),因为传统的文本资料(textual source)只能讲述人类历史的部分往事。^①但是,受限于技术,也受限于“视觉素养”(visually literate)的欠缺,^②历史学家对非文本材料,往往局限在锦上添花式的零星使用,无法对大规模的图像素材进行全方位分析,背后的原因主要是,在技术层面,数字工具对图像的识别效果还不是特别可靠。图像属于富文本资料,要充分有效地利用图像,比如历史报刊上的政治讽刺漫画,需要对漫画的丰富细节进行准确还原:包括但不限于时间、标题,并能够对画面信息进行文字描述。^③在实践环节,让电脑软件实现上述功能,并能保持足够的准确度,存在诸多棘手的技术难题。比如,在历史报纸收藏中,当数字图像是从早期的微缩胶卷创建而来时,会普遍出现所谓的“噪点效应”(noise effects)。噪点效应会对页面的主要信号带来干扰,包括不均匀分布的亮度、从页面另一侧可见的字符等,对人类识别或者计算机识别,以及随后的处理都会带来麻烦。^④不过,在人工智能的协助下,技术难题逐步被攻克,让批量识别图片具有可行性。在新技术的支撑下,大规模历史图像的数据库被构建,基于图像数据库的学术研究也将得到充分发展。

二、人工智能对齐历史学家的工作流程

从大语言模型的工作原理上看,生成式人工智能之所以具备“智能”,恰恰在于它的学习能力是向历史学家的工作方式对齐的结果。有学者直言,人工智能就是历史学,因为作为人工智能核心技术支持的机器学习,其算法原理就是对“历史”数据的分析和处理,与历史学的思考逻辑非常吻合。^⑤简言之,人工智能在像历史学家那样整理、分析和解释,实现着从过去已知获得未来新知的能力。无独有偶,独立学者约书华·斯特恩菲尔德(Joshua Sternfeld)也有类似的观点,他在一篇文章中详细讨

① Catherine Armstrong, *Using Non-Textual Sources*, Bloomsbury Academic, 2016, pp. 1-2.

② Peter Burke, *Eyewitness: The Use of Images*, Cornell University Press, 2001, p. 10.

③ Ludmilla Jordanova, *The Look of the Past: Visual and Material Evidence in Historical Practice*, Cambridge University Press, 2012, pp. 9-11.

④ Leen-Kiat Soh, Elizabeth Lorang, Chulwoo Paek, and Yi Liu, “Applying Image Analysis and Machine Learning to Historical Newspaper Collections”, *The American Historical Review*, Vol. 128, No. 3, 2023, p. 1382.

⑤ Marnie Hughes-Warrington, “Toward the Recognition of Artificial History Makers”, *History and Theory*, Vol. 61, No. 4, 2022, pp. 108-113.

论了“AI作为历史学家”(AI-as-Historian)的可能性,认为人工智能与历史学家在三个维度上极为相似:源数据的获取和选择、通过神经网络对数据进行语境化处理,以及基于反馈算法进行经验学习。^①

如果这样的类比是合适的,那么在可以预见的不远将来,历史学家的工作将被概括为人工智能驱动(AI-driven)的研究。^②在历史学经历了诸如“空间转向”“数字转向”之后,不可避免地来到“人工智能转向”的时代,这将意味着能够为历史研究提供更广阔的空间。我们以历史模拟(simulation)为例来说明人工智能对历史研究效率的提升。

历史模拟一直是部分历史学家尝试达成的设想。通过对历史的模拟,特别是对战争的模拟,让历史学家获得对历史细节的把握,“不仅能让人深入了解历史事实、日期、人物或事件,还能让人深入了解历史的偶然性、条件和环境的复杂,这些都是真正了解历史的基础”;^③在沉浸式体验历史的基础上,研究者对历史问题的分析将具备基于语境的厚重感。^④正是出于这样的愿望,有学者注意到电子游戏与历史研究之间跨界融合的可能性。在他们看来,电子游戏打造了模拟历史的数字环境,能够为学者提供进行历史实践的途径。^⑤但是,基于电子游戏的历史模拟存在很多技术性问题,比如游戏的开发需要高成本投入等,历史研究也曾经借用基于主体的建模(agent-based modeling)进行模拟,但这样的方法需要前期学习成本,达到的效果受算法影响可能有一定局限性。

在生成式人工智能技术的辅助下,历史模拟会变得更加容易,参数的调整也将更加便捷;历史学家也不需要掌握机器学习的知识背景或编写程序语言的技能,^⑥就能开展定制化的历史模拟,因此不必把精力投入历史模拟的技术环节,而只需集中讨论学术指向的研究问题。多智能体模拟(multi-agent simulation)技术将成为一个极具前景的发展方向,斯坦福大学的研究团队建立了虚拟小镇,模拟人类社区的运行状况。^⑦罗格斯大学(Rutgers University)的研究团队则将大语言模型化身为智能体(AI agent),模拟战争冲突中的不同角色,讨论引发战争的因素和条件。^⑧通过对历史事件的模拟,该项目的研究将帮助历史学家重新定义解决冲突和维护和平的战略方法。如果能够将更多要素,比如地理、街景、感知、经历等信息赋予智能体,让其在对真实世界的模拟环境中完成各种复杂任务,^⑨实现对复杂历史事件的复现,将能够更有效地回应具体的历史议题。

实际上,历史学家在自己的工作流程中引入生成式人工智能,门槛不在于技术,而在于主观意愿。数学家陶哲轩自生成式人工智能发布以来,就一直在尝试将它引入数学理论问题研究的工作流

① Joshua Sternfeld, “AI-as-Historian”, *The American Historical Review*, Vol. 128, No. 3, 2023, pp. 1372 – 1376.

② 谢天等:《大模型时代的社会科学,何去何从?》,《图书情报知识》2023年第6期。

③ Matthew Kapell and Andrew Elliott, eds., *Playing with the Past: Digital Games and the Simulation of History*, Bloomsbury, 2013, p. 13.

④ Philip Sabin, *Simulating War: Studying Conflict through Simulation Games*, Continuum, 2012; William Illsley, “Digital Surrogacy: Politics and Aesthetics in Visualising the Historical Past of a City”, *International Journal of Heritage Studies*, Vol. 28, No. 2, 2022, pp. 216 – 234.

⑤ Adam Chapman, *Digital Games as History: How Videogames Represent the Past and Offer Access to Historical Practice*, Routledge, 2016.

⑥ Philippe Giabbanelli, “GPT-based Models Meet Simulation: How to Efficiently Use Large-scale Pre-trained Language Models across Simulation Tasks”, arXiv, 21 Jun. 2023, <https://arxiv.org/abs/2306.13679> [2023 – 12 – 25]

⑦ Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, et al., “Generative Agents: Interactive Simulacra of Human Behavior”, arXiv, 6 Aug. 2023, <https://arxiv.org/pdf/2304.03442.pdf> [2023 – 12 – 20]

⑧ Wenyue Hua et al., “War and Peace (WarAgent): Large Language Model-based Multi-Agent Simulation of WorldWars”, arXiv, 30 Jan. 2024, <https://arxiv.org/abs/2311.17227> [2024 – 03 – 20]

⑨ Jihan Yang et al., “V-IRL: Grounding Virtual Intelligence in Real Life”, arXiv, 5 Feb. 2024, <https://arxiv.org/abs/2402.03310> [2024 – 03 – 25]

程。他在 2023 年的一篇学术论文中,就披露了用 GPT 辅助证明不等式定理的过程。^① 作为与历史研究实践更具适配性的生成式人工智能,它进入历史学家的工具箱更是水到渠成的事情。历史学家不用担忧学习新工具的时间成本,因为人工智能发展的初衷就是让用户的操作更简捷、更方便,用“奥卡姆的剃刀”尽可能地扫除冗余障碍。^② 历史学家需要做的仅仅是用专业的学术语言提出问题,而这是历史学家工作流程中的基本操作。

总而言之,历史学家如果能够善用生成式人工智能进行历史模拟的工作,将会在学者与“智能体”之间形成良好的学术互动,一方面能够对生成式结果进行专业评估,另一方面也实现了人工智能的工作流程向人类工作对齐,极大降低出现“幻觉”(hallucination)的可能性。

三、人工智能会取代历史学家吗?

当然,有人可能会担心,虽然人工智能的工作原理是对历史学家的模仿,但是,它在算力和存储能力上远超人类,难道不会对历史学家带来威胁吗?笔者依然乐观地估计,这种担忧在目前看来是多虑了。

首先,生成式人工智能还没有智能到可以自主开启一个研究问题,更别说开拓全新的研究领域了。在大语言模型的“加持”下,历史学家的研究能力和工作效率确实能够得到大幅度提升,不仅一个人就可以组建一个研究团队,而且乐观的学者都在展望更加令人心动的人机共生的未来图景:历史学家也能够经由提升人工智能素养(AI literacy)达到智能增强(intelligence augmentation)的目的,最终不仅能够消弭人机冲突,还能实现人机共生的协助关系。换句话说,学者们追求的是以人为中心的人工智能(human-centered AI),^③最终实现的是人机协同工作,而不是机器取代历史学家。实际上,早在 20 世纪 90 年代,就有学者提出“智能伙伴”(intelligent partnerships)的概念,认为人与智能技术之间的协作可以超越单独个体的智能。^④ 当然,从目前历史学家的工作方式来看,人工智能还不足以主动开启研究议题,遑论取代历史学家了。即便将来出现更先进的人工智能算法,“历史学家在回路中”(historian-in-the-loop)也是不能放弃的必然选项,^⑤只有这样才能确保研究议题的学术性,减少偏见,纳入稀缺史料,增加透明度。

此时的历史学家,不会再有“独狼式”的孤独感,因为他可以在人工智能的帮助下实现不同的分身,完成不同领域的工作。历史学家的工作节奏,将越来越摆脱传统的孤身作战,而是有统领团队的气势。这也意味着,比起个人事必躬亲,历史学家更需要关注各项工作的协调,需要更多务虚的思考,为研究问题注入更多跨学科的知识背景。此时,研究问题的设置就显得更加重要了,而且要向大

① <https://mp.weixin.qq.com/s/Cezn03Xksu4XLiQYFE9cug>[2024-03-25]

② “奥卡姆的剃刀”是中世纪哲学家奥卡姆提出的逻辑原则,主要宗旨是强调非必要,勿增实体。这个比喻常常用来鼓励人们在解决问题时,寻找更直接、更简洁的方案。

③ Mark Riedl, “Human-centered Artificial Intelligence and Machine Learning”, *Human Behavior and Emerging Technologies*, Vol. 1, No. 1, 2019, pp. 33-36.

④ Eva van Dis, Johan Bollen, and Claudi Bockting, “ChatGPT: Five Priorities for Research”, *Nature*, Vol. 614, No. 7947, 2023, p. 226.

⑤ 此处借用了相关学者提出的“人在回路中”(Human-in-the-loop),参见 Robert Munro, *Human-in-the-loop Machine Learning*, Manning Publications, 2021。

语言模型清楚交待研究问题的由来、方向、展开等细节,往往需要与 GPT 进行多轮沟通,就像合作伙伴那样,清楚表明自己的需求、预想及目的。“历史学家在回路中”的工作流程,意味着学者始终是掌控者,只有在这种方式下,对 GPT 的使用才是有效且有价值的。

其次,历史研究的原创性来源之一是新史料的使用,这也是兰克史学强调档案材料的原因。至少在目前,大语言模型的智能得益于强大算力的支持,但不论它有多么渊博的知识,其认知来源都是训练素材。简言之,它的知识边界就是文献数字化的边界。没有被数字化的材料,再强大的大语言模型也一无所知。我们不要天真地以为,所有的历史文献都完成了数字化,现实情况是,网络时代普遍存在的数字鸿沟现象,在大语言模型时代依然存在;实际上,在各个档案馆里,还“沉睡”着大量原生态的文献,这些文献不仅没有被数字化,更不用说作为数据集被提供给大语言模型进行训练了。

让我们回到之前“讽刺漫画”的例子。虽然大语言模型给出了非常详细的解读和准确的出处,但这并不是因为它无所不知,而是恰好这幅漫画在互联网平台很容易被检索到,德国历史文献在线数据库“文献与图像中的德国历史”(Deutsche Geschichte in Dokumenten und Bildern)收录了这幅漫画,并对它的内容、背景、寓意等有详细描述。^①大语言模型仅仅借助其强大的检索功能,便能找到已有的信息来源。实际上,如果换一张在网上找不到相关信息的图片,大语言模型的答案将漏洞百出。笔者也做过相关试验,效果确实不尽如人意。

简言之,目前的生成式人工智能还无法实现基于新史料的历史书写。如果人工智能不能吸纳新材料进行内容生成,那么它的历史写作将只能囿于老生常谈、人云亦云,既不能兑现好的历史书写需立足于扎实文献的基本要求,也无法形成独到的洞见,完成原创性研究。毫无疑问,历史研究的任务就是要深入挖掘、分析和解释丰富的历史文献,原创性研究来自对原始文献的解读,获得对过往事件和人物的准确理解,并找出其中的因果关联,形成新的见解和观点。大语言模型如果接触不到新史料,其历史研究的原创性就是无源之水、无本之木。

当然,我们可以在研究过程中通过拍照、手工录入等方式完成对原始文献的数字化,然后投喂给大语言模型进行个性化训练,让它帮助历史学家进行整理、分析和解读,但在这个工作流程中,主动性显然掌握在历史学家手中,大语言模型的被动性一如既往。如果没有历史学家先期开展档案文献整理工作,后续一切流程都无法开始。因此,生成式人工智能还需要跟历史学家密切配合,才能贡献具有学术意义的内容;否则,生成的文本只能是敷衍塞责的八股文章,不仅缺乏创新,内容的准确性也会大打折扣。

最后,历史研究的内容不仅有文本的生产,还有基础性的史料整理工作。历史研究的工作基础是史料,对历史文献进行整理,形成了历史研究的基础设施。但是,文献整理是典型的实操工作,需要历史学家将大量的时间和精力投入查阅档案馆、田野调查、口述访谈等工作中,这些都是生成式人工智能目前无法独立完成的。

这是当下生成式人工智能最大的短板,因为大语言模型缺乏行动力,不能在物理世界有效展开活动。当然,不排除在历史学家整理史料的过程中,在特定的环节使用大语言模型作为辅助。例如,在进行口述访谈时,语音识别和自动翻译技术可以帮助历史学家快速准确地转录和翻译采访内容,大大提

^① https://ghdi.ghi-dc.org/sub_image.cfm?image_id=1326&language=german[2023-12-16]

升研究的效率和精度;人工智能的信息检索技术,能够帮助历史学家迅速提取文献,发现文献中的模式和趋势。不过,我们要当心,在文献梳理的过程中,大语言模型可能出现“幻觉”,导致其生成的结果不是那么令人信服。有学者专门做过定量分析,结果发现,用大语言模型进行某个研究主题的文献梳理,可能会有 66% 的文献是杜撰出来的。^① 这进一步说明了大语言模型在文献整理方面的无能为力。

四、人工智能参与历史研究实战

即便最热烈鼓吹生成式人工智能的学者也承认,对算法产生的结果要保持警惕。那么问题来了,使用者去监控人工智能生成的结果是否正确,是增加了历史学家的负担,还是减轻了他们的工作量?在笔者看来,生成式人工智能的结果,在专业学者眼中,并不意味着比其他形态的史料更真实,也不一定更令人怀疑;对证据采取批判性分析的态度,本是历史学的学科属性决定的,也是历史学家的分内之事。笔者想用一個具体案例来进一步回应这样的问题。

有一位朋友发来一张邮票的图片,让笔者查询一下邮票上的人物,以及邮票的发行信息。笔者本身对邮票没有研究,好在邮票上写出了人物的姓名,应该是德意志帝国时期著名的思想家莱辛。不过,由于他的名字使用了简写,而且有一部分字母被邮戳覆盖了,并不能一眼看出来。笔者突发奇想,ChatGPT 识别真实图像史料的表现力究竟有多强?

说干就干,笔者把这张图片甩给了 ChatGPT,首先给它提出的问题是,邮票上的人物是谁?结果并不在笔者的意料之外,ChatGPT 把人物识别成了歌德,还煞有介事地向笔者简短介绍了歌德的代表作。笔者当然不满意,但也觉得情有可原。启蒙时代的人物肖像画很容易搞混,而且 ChatGPT 能够定位到“歌德”,应该是在开盲盒式的竞猜中,大概率最正确的结果。笔者同时在 Gemini 中测试了一下,这个最新的大语言模型一开始也给出了歌德的答案。这比较符合大语言模型基于概率的预期,也似乎能够说明两个大语言模型在训练集的组成上有非常相似的来源。

笔者决定再给 ChatGPT 一次机会。专家都有看走眼的时候,何况人工智能呢?于是,笔者直接指出它的答案是错误的,让它再仔细“想想”。ChatGPT 是一位知错能改的助手,谦卑地承认了错误,“思索”了一会后,又给出了莱布尼茨的名字。只能说,真相又接近了一步,这个答案表明,ChatGPT 看到了印在邮票上的缩写“Goth”,但是没有正确识读。

孺子可教也。接下来,笔者给出新一轮提示,让 ChatGPT 重点关注邮票左侧的字母。结果,非常令人惊艳的是,在笔者的循循善诱下,它终于识别出邮票上的人物是德意志帝国时期著名的思想家莱辛。

这还不够,笔者想知道这张邮票是什么时候发行的。当这个问题抛给 ChatGPT 后,它开始自主在网络上搜索相关信息,笔者注意到它使用的关键词包括“lessing german stamp”,得到的结果,不仅告诉笔者发行日期是 1961 年,还给出了一个 URL 地址,点击进入,是维基百科上的一张邮票,虽然票面是莱辛的肖像画,但显然不是我们要找的那张。

ChatGPT 忽略了邮票上更多的线索。在这张邮票的底部有一行字,“deutsches reich”(德意志帝

^① Michael Haman and Milan Skolnik, “Using ChatGPT to conduct a literature review”, *Accountability in Research*, published online: 6 Mar. 2023, <https://doi.org/10.1080/08989621.2023.2185514>[2023-12-20]

国),用花体字印刷,并不好识别。稍微有一点德国历史知识背景的读者都能推断出,这张邮票不可能在德意志帝国早已瓦解的1961年发行。

笔者把这个额外的信息分享给 ChatGPT,它开始了新一轮搜索。此时,新添加进去的搜索关键词包括了“deutsches reich”。最终,它给出了正确的年份1926年,并附上了一个URL链接,点击进去,正是跟我们手头一模一样的邮票。

总结起来,这是一次令人惊艳、具有实战色彩的使用体验:一方面验证了 ChatGPT 的有效性,另一方面更加明确地显示出,ChatGPT 的正确打开方式,需要用户将其视为能够对话的研究助理,通过不断尝试与交流,最终获得满意的结果;作为用户,我们不仅要能够辨识结果的可信度,还需要不断给 ChatGPT 提供新的线索。很多人工智能的业内人士强调,取代程序员工作的不是人工智能本身,而是那些善于与人工智能协同工作的程序员;同理,生成式人工智能也无法取代历史学家,但是,擅长将人工智能纳入研究流程的历史学家将极大地冲击那些还不关注人工智能的同行。虽然人工智能之于史学界还不至于是一个颠覆性的“搅局者”,但它带来的改变将逐渐显现。

结 语

人工智能的发展,对人类社会带来的挑战是持续而全面的,学者们很早就开始关注这个议题。2019年在“数字人文主义维也纳宣言”(Vienna Manifesto on Digital Humanism)中,相关学者注意到人工智能“双刃剑”的问题:在自动决策和人工智能的时代,创造力和对人性的关注至关重要。^①从数字人文(digital humanities)到数字人文主义(digital humanism)的转换中,学者强调无所不在的数字化、智能化首要的考虑是满足人的价值和需求,技术是为人服务的,而不是相反;我们需要把“人类”放在工作的中心位置。

要想兑现人工智能的发展以服务人类为中心的承诺,可以在技术层面为人工智能的发展加装“围栏”。比如,数字史学专家在论及未来发展趋势时,都会强调数据的标准、算法的透明,^②以此让数字史学的方法具有可验证性与可信度。数据的标准化对于确保研究结果的准确性至关重要,因此,数字史学专家强调采用一致的数据标准处理史料,以便在研究中进行比较和分析。这些标准包括数据的收集方法、处理过程及验证实践等。此外,算法的透明度也不可或缺。专家期望通过了解算法的具体原理和操作方式,理解数字技术生成结果的机制。仅依赖结果无法判断研究的准确性,算法的透明将有助于同行专家对结果的可靠性进行审查。

生成式人工智能接入历史研究之后,同样的期待和呼吁依然适用:大语言模型的训练数据集、算法原则等都需要得到学术共同体的认可,才能提升人工智能辅助历史研究的效果和程度。这意味着,数据集的选择和处理都应该符合学术规范,并具有广泛的代表性。在选择训练数据时,应注意不偏袒某一特定群体或地域,以避免引入偏见,加剧数字鸿沟。此外,算法原则也需要得到学术共同体的认可,以确保其可靠性和可行性。算法的设计和操作应当经过严谨的审查,具有透明性,以便对结果进行验证。

^① Hannes Werthner et al., *Perspectives on Digital Humanism*, Springer, 2022, p. xiii.

^② Anna Siebold and Matteo Valleriani, “Digital Perspectives in History”, *Histories*, Vol. 2, No. 2, 2022, pp. 175 – 176.

人工智能驱动的历史研究,恰如数码相机时代的绘画艺术一样,并非零和博弈。从写实的角度看,绘画技艺显然敌不过哪怕最低端的数码相机;但是,在艺术创意的维度上,绘画艺术依然有极强的生命力。历史研究同样如此。人工智能“加持”的历史学家的确能够实现智能增强,但因循传统路径的学者也并不会走投无路,两者可以在不同维度上找到自己的定位。

这种乐观主义的底气来自于人文主义的价值,也正是包括历史学家在内的人文主义者的核心竞争力所在。实际上,人工智能尽管是当下最尖端的科技,但它依然无法撼动历史研究的一些基本要素。英国历史学家爱德华·霍列特·卡尔(Edward Hallett Carr)在反思历史研究客观性时,提醒同行要有更广阔的视野。按照卡尔的说法,客观的历史学家不是单纯地寻找事实,^①还意味着“有能力将自己的视野投射到未来,从而对过去有更深刻、更持久的洞察力,这一点是那些将视野完全局限于自己眼前境况的历史学家无法企及的”。^②在卡尔写作的时代,数字技术并没有进入人文学者的日常,人工智能被应用于历史研究更是闻所未闻,但在他为历史学家预设的宏伟气象中,要有“朝向理解未来”的目标,人工智能为历史研究带来了宏大视角,历史学家也要学会驯服人工智能这个“大杀器”,^③强有力的手段就是人文性。

从理论倡导到研究实践:数字史学未来走向刍议

梁 晨 (南京大学历史学院教授)

近年来,数字技术的迅猛发展催生出一系列数字化分析工具——各类文本、关系网络和地理信息分析工具不断出新,能自主挖掘材料并形成具有创意文本或视频的大语言模型、文生视频模型更是横空出世,数字技术似乎已能介入从史料收集、材料分析到文本写作和史实再现等全过程的史学研究环节。史学界对依靠数字技术推进历史研究的期待与日俱增,学术热度也愈发凸显:仅2022年,“人工智能、大数据与历史学融合发展”入选中国历史研究院评选的“中国历史学十大研究热点”、“数字人文”正式成为教育部批准的本科专业、北京大学历史学系开始招收“人工智能与历史学”方向的博士研究生。

不过,在目前中国数字史学甚至数字人文学界,学理层面的分析和对未来的展望较多,而基础设施建设和实证研究均很不足。整个领域给人以理论倡导者多、付诸实践者少的感受。与此同时,许多历史学者对数字技术是否真地能够有助于历史学研究心存疑虑。例如,成一农认为,历史学的终极目标是有目的、有人性地“解释”和“预测未来”,数字技术对此没有多少帮助。^④不管是否认同成一农的看法,总体而言,史学界对数字史学尚持观望之态,很少有研究者下场一试是不争的事实。一份2023年的中国数字人文学术体系调查研究指出,当前中国数字人文存在的首要问题是“学科发展

① Richard Evans, *In Defence of History*, Granta Books, 2000, p. 225.

② Edward Carr, *What is History?*, Penguin Books, 1987, p. 123.

③ “大杀器”当然只是一个比喻,但人工智能具有潜在的风险是有目共睹的事实。人工智能的训练需要消耗大量化石燃料,带来了更多温室效应的排放,破坏了环境;如果人工智能被别有用心的人用来开发核武器、生物武器,对于人类世界则是切实的威胁。所以,驯服人工智能,不是杞人忧天,而是非常迫切的需求,参见 Guglielmo Tamburrini, “Artificial Intelligence and Large-Scale Threats to Humanity”, in Hannes Werthner et al., *Introduction to Digital Humanism*, Springer, 2024, pp. 241 – 254.

④ 成一农:《抛弃人性的历史学没有存在价值——“大数据”“数字人文”以及历史地理信息系统在历史研究中的价值》,《清华大学学报》2021年第1期。