

AI 时代历史学的价值升华与角色重塑

——一个基于前馈神经网络模型的理论阐释

张光伟

(陕西师范大学 历史文化学院, 陕西 西安 710119)

摘要: 人工智能 (AI) 的飞速发展正深刻重塑知识生产的全景, 史学研究亦面临范式革新的重大机遇与挑战。当前学界对 AI 的应用多停留于效率工具的层面, 忽视了生成式 AI 深度融合与自主学习能力可能带来的学科反噬风险。文章旨在超越工具论, 探讨历史学如何与 AI 构建可持续的双向赋能关系。我们认为, AI 时代历史学独特的跨学科整合能力并非简单的信息叠加, 而是一种价值升华机制。借鉴前馈神经网络模型, 文章将历史学家的深度思维阐释为一种独特的元认知中间层, 它通过赋予意义、提供时间智慧、注入人文关怀和激发反思性批判, 实现对跨学科输入的价值再造。AI 不仅能通过跨语言文献解析等手段极大增强该网络的输入层, 更能借助思维链等技术优化其中间层的运行, 创新其输出层的形态。文章认为, AI 时代历史学家的核心价值将从知识记忆与处理, 转向提出深刻问题、主导人机协同、进行价值判断与意义建构。这不仅是历史学应对技术挑战的生存之道, 更是其在新的知识生产体系中重塑核心地位、贡献独特智慧的必由之路。

关键词: 人工智能; 跨学科整合; 价值升华; 前馈神经网络; 史学方法论

中图分类号: K01; TP18 **文献标识码:** A **文章编号:** 0257-0246 (2026) 01-0131-08

引言

21 世纪 20 年代以来, 以大语言模型 (LLM) 为代表的生成式人工智能 (AI) 技术取得了突破性进展, 在全球范围内引发了一场深刻的社会变革。正如 AI 领域的领军人物李飞飞教授在 2025 年巴黎 AI 峰会上所预言的, 未来的历史学家无疑会将我们所处的时代称为“第一个真正的 AI 时代”, 其影响可能具有“文明级别的后果”。学术界包括人文社会科学领域, 正以前所未有的热情拥抱 AI 技术。环顾国内高校与研究机构, 争相部署 AI 大模型、推出各类与 AI 结合的业务系统已成普遍现象。然而, 在这股涌动的热潮之下, 理论与实践的迷思也随之浮现: 当前对 AI 的应用, 多是将其视为继信息技术、互联网技术之后的又一个效率提升工具, 沿用旧有的技术应用模式, 而缺乏对 AI 本质, 尤其是其强大的自主学习、推理与深度融合能力的深刻理解。

这种将 AI 简单工具化的认知, 潜藏着巨大的风险。我们必须认识到, 生成式 AI 与以往任何技术都存在本质差异。它不再仅仅是作用于研究客体的工具, 而是能够深度介入乃至模拟研究主体核心认

基金项目: 教育部哲学社会科学研究重大课题攻关项目 (22JZD037)。

作者简介: 张光伟, 陕西师范大学历史文化学院讲师, 研究方向: “人工智能+历史学”交叉学科。

知过程的伙伴^①或对手。如果一个学科领域与 AI 无法形成持续的、双向的价值互动闭环，即该领域无法为不断进化的 AI 提供新的、有价值的输入（包括独特的知识、数据、问题乃至方法论等），那么这个领域可能面临被 AI 反噬的风险。这并非危言耸听，它表现为：当学者仅仅单向地利用 AI 完成工作，而自身核心能力不随之升级，AI 的强大自主学习能力将使其迅速超越人类；最先出现的可能是这样一种现象，传统学者不是被 AI 直接取代，而是被掌握了 AI 并能与之高效协作的新一代学者所取代^②。

面对 AI 的强大学习能力和潜在的反噬风险，历史学是否注定会走向衰落？或者，历史学是否具备某些独有的特质，使其能够与 AI 构建一种可持续的双向赋能关系？这正是本文的核心关切。我们认为，历史学作为一门古老而智慧的学科，其独特的跨学科整合与价值升华能力，正是构建这种良性关系的关键。

AI 所带来的深刻变革，不仅仅是技术层面的，更是认识论和方法论层面的，这早已引起了历史学界的高度关注。美国历史学会（AHA）旗舰期刊《美国历史评论》（*The American Historical Review*, AHR）2023 年推出的“人工智能与史学实践”专题，系统性地探讨了 AI 对历史学核心概念的冲击^③。学者们普遍认识到，AI 不仅是一个工具，它正以前所未有的深度，介入到历史意识的形成之中，甚至展现出某种历史能动性^④。然而，这场讨论也暴露了一个核心困境：在承认 AI 有巨大潜力的同时，历史学界对其黑箱式的运作、潜在的偏见，以及对严谨研究方法的冲击，充满了深深的忧虑^⑤。AHA 在 2025 年 8 月 5 日发布的《历史教育中的人工智能指导原则》^⑥中，更是将“批判性地评估 AI 的论证过程”作为核心素养来强调。那么，我们能否创造一种新范式，既能利用 AI 强大的计算能力，又能确保其整个工作流程是透明的、可控的，并且严格遵循历史学的核心方法论与学术伦理？

本文旨在超越工具论，深入探讨 AI 时代历史学独特的学科价值、其在跨学科研究中的核心机制，以及 AI 作为催化剂对历史学家角色的重塑，从而为历史学在新的知识生产体系中寻找生存与发展之道。

一、跨学科整合的时代呼唤与历史学角色的重估

AI 时代是一个内在地强烈呼唤跨学科整合的时代。这并非一句口号，而是源于我们所面临问题的性质以及知识生产模式的根本性变革。当代社会面临的诸多重大挑战，无论是气候变化、公共卫生安全，还是人工智能伦理与治理，都呈现出典型的棘手问题（Wicked Problems）特征：它们边界模糊、成因复杂、利益相关方多元，且没有一劳永逸的简单解决方案^⑦。应对这些挑战，单一学科的知识和方法已经捉襟见肘，全球学术界正积极倡导一种知识融合的新范式，即通过深度整合不同学科的知识、方法和思维模式，来催生颠覆性的科学发现，并解决复杂社会问题^⑧。这种范式要求我们必须打破传统学科壁垒，以问题导向取代学科导向。

① Zhang Guangwei, "How AI is Becoming the Perfect Partner for Historians," <https://www.sixthtone.com/news/1017267>.

② K. Lakhani and A. Ignatius, "AI Won't Replace Humans—But Humans with AI will Replace Humans without AI," *Harvard Business Review*, <https://hbr.org/2023/08/ai-wont-replacehumans-but-humans-with-ai-will-replace-humans-withoutai>, August 4, 2023.

③ R. Darrell Meadows and Joshua Sternfeld, "Artificial Intelligence and the Practice of History: A Forum," *The American Historical Review*, Vol. 128, No. 3, September 2023, pp. 1345–1349.

④ Joshua Sternfeld, "AI-as-Historian," *The American Historical Review*, Vol. 128, No. 3, September 2023, pp. 1372–1377.

⑤ Kate Crawford, "Archeologies of Datasets," *The American Historical Review*, Vol. 128, No. 3, September 2023, pp. 1368–1371.

⑥ American Historical Association, "Guiding Principles for Artificial Intelligence in History Education," <https://www.historians.org/resource/guiding-principles-for-artificial-intelligence-in-history-education/>.

⑦ Horst W. J. Rittel, Melvin M. Webber, "Dilemmas in a General Theory of Planning," *Policy Sciences*, Vol. 4, No. 2, June 1973, pp. 155–169.

⑧ 美国科学院研究理事会：《会聚观：推动跨学科融合：生命科学与物质科学和工程学等学科的跨界》，王小理、熊燕、于建荣译，北京：科学出版社，2015 年，第 1 页。

中国高等教育提出的“四新”建设，特别是新文科建设，正是对这一时代呼唤的深刻回应与战略布局。新文科建设的核心要义，在于推动文科教育的交叉融合与创新发展，使其能够更好地应对新技术革命带来的挑战，并为解决复杂的当代中国与世界问题提供深厚的智慧支持^①。它强调文科与理工农医的深度交叉，培养能够进行批判性思维、复杂问题求解和跨文化沟通的复合型人才^②。

在这一背景下，重新审视历史学在跨学科整合，特别是在新文科建设中的角色与价值，显得尤为迫切。传统上，历史学在跨学科研究中常被定位为背景提供者、案例素材库或人文点缀者。这些认知并非错误，但其被动性与辅助性定位极大地窄化了历史学的价值，忽视了其在思维方式、价值判断和意义建构上的独特贡献。事实上，历史学的视角与方法对理解任何知识领域都具有根本性价值。一个显著的现象是，从物理学、化学、医学等自然科学，到经济学、社会学等社会科学，几乎所有成熟的学科都内生出对其自身发展历程进行历史性考察的需求，并形成了各自的学科史。这雄辩地证明，历史是理解任何知识体系的元框架，历史学的思维方式具有某种底层通用性。

此外，历史学作为现代学术研究的重要奠基者，其发展塑造了诸多基本学术规范，如严谨的参考文献与注释规范、批判性的史料辨伪精神等，这些已成为所有学科共同的方法论遗产。然而，历史学更核心的方法论精髓——例如对时间性、过程性、情境性、复杂因果性的深刻理解，以及从碎片化信息中构建连贯叙事的能力——对其他学科的影响往往是潜移默化的。AI 的出现为这些“内功心法”的显式化提供了历史性契机。AI 技术，尤其是 LLM，其内在运行机制与历史学关注的过程、情境、关联和叙事不谋而合，这使得历史学那些深刻的方法论智慧，有了被更清晰地识别、更系统地应用、更有效地赋能其他学科的技术基础。因此，我们亟须超越传统认知，对历史学在 AI 驱动的跨学科整合中的核心价值进行深度挖掘。

二、历史学在 AI 时代的跨学科价值机制

历史学在 AI 驱动的跨学科研究中的核心竞争力，并非简单的信息整合，而是一种可称为“价值升华”的独特机制。它如同点金之手，能够将来自不同学科的、相对孤立的材料——无论是经济学模型还是地理学数据——进行深度加工和意义重塑，催生出“1+1>2”的质变效应。在 LLM 能够轻易完成信息整合的今天，这种机制的价值尤为凸显。它主要通过以下四个相互关联的层面实现：

1. 建构解释性框架，对抗 AI 的事实碎片

很多学科，乃至 AI 本身，能够提供海量的点状事实，但这些点在未经组织时是零散且缺乏意义的，或者是自说自话的。历史学的首要价值升华机制，便是通过其核心的叙事能力，将这些点编制进一个具有内在逻辑和深层解释力的“意义之网”中。这并非简单的故事讲述，而是如海登·怀特所揭示的，一种通过“情节化”（Emplotment）来主动建构意义的深刻认知活动。^③ 一个 AI 可以检索出一千条关于某次经济衰退的数据，但只有历史学家能够将它们置于一个王朝末路的悲剧叙事或改革阵痛的希望叙事中，从而揭示其中不同的历史意蕴。这种从事实到意义的“跃迁”，正是对抗 AI 生成内容有知识而无见识这一弊病的根本武器。

2. 洞察动态过程，超越 AI 的静态快照

许多学科及当前 AI 模型处理数据的方式，本质上是共时性的、切片式的分析。它们擅长处理此时此刻的状态，却难以理解从过去到现在的复杂过程。历史学则为跨学科研究注入了不可或缺的时间

^① 樊丽明：《新文科建设：走深走实 行稳致远》，<https://www.chsi.com.cn/jyxx/202105/20210510/2066701201.html>。

^② 教育部：《新文科建设工作会在山东大学召开》，http://www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/202011/t20201103_498067.html。

^③ Hayden White, *Metahistory: The Historical Imagination in Nineteenth-Century Europe*, Baltimore: Johns Hopkins University Press, 1973; 彭刚：《叙事、虚构与历史——海登·怀特与当代西方历史哲学的转型》，《历史研究》2006年第3期。

智慧。通过引入历时性视角、长时段观念与过程性思维，历史学关注变化与持续性、断裂与延续性。这种将静态数据快照还原为动态演化进程，并洞察其内在逻辑的能力，是 AI 单纯依靠数据模式识别所难以企及的。它为理解任何复杂系统（无论生态、经济还是社会）的演化，提供了根本性的认知框架。

3. 锚定人的经验，校准 AI 的技术理性

偏重技术、模型或宏观结构的跨学科研究，以及 AI 系统内在的技术理性，很容易忽略个体经验、情感需求、文化价值和伦理维度。历史学始终将“人”及其复杂的经验世界置于历史进程的中心。它关注人的能动性、多样性经验、价值观念的变迁以及社会公平正义问题。这种能力，可以将 AI 生成的某地区 GDP 变化的冰冷分析，与该地区个体的命运沉浮、悲欢离合联系起来，从而为研究赋予温度和灵魂；它更是在 AI 时代为纯粹的技术和数据分析，提供价值参照和伦理校准的关键机制，从而实现从物本逻辑向人本逻辑的回归。

4. 审视知识前提，打破 AI 的范式局限

每个学科，乃至每个 AI 模型，都有其不言自明的理论假设和范式边界。历史学的最高阶价值，可能在于通过对其他学科知识体系（理论、概念、算法）进行历史化、情境化的考察，来激发学科对话中的元认知反思。它挑战“想当然”的前提，揭示某个模型或理论背后隐藏的偏见和未被审视的假设。正如我们在若干年的历史学与技术等多学科融合实践中的深刻体验：历史学的融入能够有效促进协作学科之间的对话，促使研究回归到更根本的历史情境与人的活动中。在 AI 时代，这种能力尤为重要，因为这种从应用知识到反思知识本身的推动力，是价值升华的最高体现。

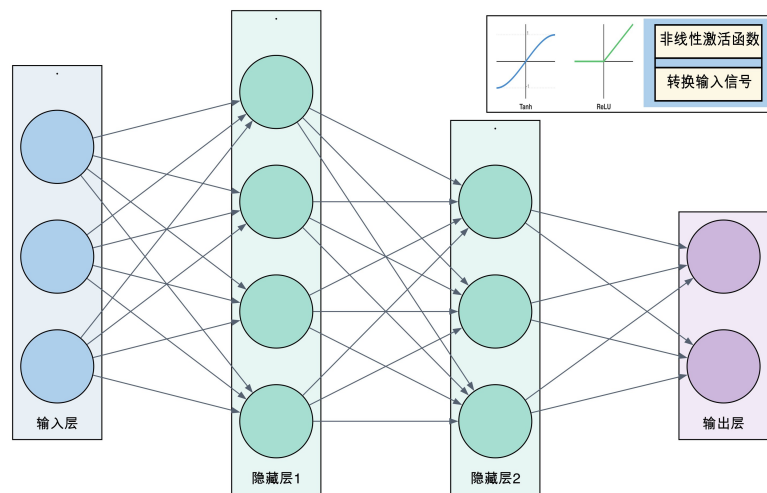
总而言之，这四重价值升华机制共同构成了历史学在 AI 时代的跨学科整合中的核心能力。可以说，它深刻回应了《美国历史评论》专题中提出的一个核心议题：面对 AI 带来的挑战，历史学家独特的价值正在于这种超越数据罗列，进行情境化、批判性意义构建的能力^①。历史学为 AI 驱动的跨学科研究提供的不仅是过去的背景，更是思考的框架和价值的罗盘。

三、作为“元认知中间层”的历史学

那么，前述的独特“价值升华”能力究竟是如何在历史学家的认知活动中协同运作的呢？为了更清晰地揭示其内在的作用机制，我们不妨借鉴认知科学与人工智能的理论模型，将历史学家的深度思维阐释为一个“前馈神经网络”模型^②。这个模型为我们理解历史学在跨学科知识生产中的独特位置和作用机制，提供了一个强大而直观的分析框架。

^① R. Darrell Meadows and Joshua Sternfeld, "Artificial Intelligence and the Practice of History: A Forum," *The American Historical Review*, Vol. 128, No. 3, September 2023, pp. 1345-1349.

^② “前馈神经网络”（Feedforward Neural Network, FNN）是深度学习的基础架构。其信息流从输入层单向传递至输出层，中间经过一层或多层“隐藏层”进行非线性处理。隐藏层通过学习数据中的模式，将低层次的输入特征逐层抽象和组合，最终形成高层次的输出。它是当前人工智能的技术起点之一。本文借用此模型，将历史学家的认知过程隐喻为一个“神经网络”，其中“隐藏层”对应历史学家将零散史料（输入）通过情境化、叙事建构、批判性反思等复杂的深度思维，升华为具有解释力的历史洞见（输出）的核心环节。



经典的前馈神经网络模型示意图

如上图所示，一个经典的前馈神经网络由输入层、隐藏层和输出层构成，图中呈现的是一个层数很少的浅层网络，而一般深度学习中的神经网络层数可达几十甚至上百层。在这个类比模型中，输入层（Inputs）对应于历史学家接触到的所有原始信息，包括海量的历史文献、考古发现、口述史料，以及在跨学科合作中来自地理学、经济学、生态学等其他学科的数据、模型和理论。输出层（Outputs）则对应历史研究最终产出的成果，如对复杂历史现象的深刻理解、对当代问题的历史洞察、具有解释力的历史叙事，以及在跨学科研究中超越单一学科局限的、更全面且更深刻的解决方案或认知图景。中间层/隐藏层（Hidden Layers）是该模型的核心，也是历史学价值升华机制发生的场所。它对应着历史学家经过长期专业训练而形成的深层认知结构与思维模式，我们称之为“元认知中间层”。它并非单一的线性处理器，而是一个由多重思维层次构成的深度网络。

然而，这个模型的真正精髓在于中间层的非线性激活函数。在神经网络中，正是这些激活函数打破了简单的线性叠加，使得模型能够学习复杂的模式，实现从量变到质变的飞跃。我们认为，前文所述的四种价值升华机制——赋予意义之网、提供时间智慧、注入人文关怀、激发反思性批判——正是历史学家头脑中那一系列复杂且强大的认知激活函数。它们对输入的“数据点”进行非线性的深加工：意义之网如同一个情节化激活函数，将事实的线性序列重组为具有整体意义的叙事结构；时间智慧如同一个注意力激活函数，在连续的时间流中非线性地凸显关键转折点的重要性；人文关怀则是一个价值激活函数，将冰冷的统计数据与个体命运相关联，激活为共鸣与伦理反思；而反思性批判无疑是最高阶的激活函数，能够跳出当前的计算框架，对整个问题的前提进行升维打击。这个过程如同中医药方中的“君臣佐使”配伍与炮制过程，对输入的“药材”进行复杂的、非线性的深度加工，从而产生远超原始输入的疗效。

这个类比模型的关键启示在于，历史学家的核心价值乃至一个学科的真正壁垒，不在于掌握了多少输入信息，而在于其中间层的深度、神经元之间复杂的连接模式及其激活函数的复杂性与精妙性。一个拥有强大元认知中间层的历史学家，能够触类旁通、点石成金；反之，则即便坐拥书城，也无法产出深刻的思想。在 AI 时代，这一特征尤为凸显，因为 AI 可以前所未有地丰富输入层的信息洪流，但设计、优化和运用这个中间层的核心激活逻辑，赋予其人文精神和批判灵魂的，依然是历史学家本人。

四、AI 赋能催化历史研究的元认知神经网络

既然历史学家的核心价值在于其独特的元认知中间层，那么我们与 AI 的协同就绝不应停留在将

其当作简单的效率工具上。为了系统性地探索一种更深层、更具共生性的人机协同范式，笔者近期在一篇专门的技术论文中提出了“知识协议工程”（Knowledge Protocol Engineering, KPE）^①这一新框架。KPE的核心是将历史学家的研究方法论编码为AI可执行的知识协议，从而对其推理过程进行引导和约束。笔者认为，KPE并非要取代历史学家的中间层，而是为前文所述的神经网络模型的每一个环节赋能，使其更高效、更强大、更具创造力。后续提到的我们正在或已经完成的相关工作，基本都是在KPE框架之下实现或重构的。KPE的理论框架在该论文中已有详细阐述，本文不做深入介绍。

1. 革命性地增强输入层

AI首先极大地拓展了输入层的边界和效率，将过去难以获取或处理的多模态材料转化为可供分析的数据。

一是海量文本的深度利用。以深度学习为基础的OCR技术，使我们能够将数以百万计的古籍和档案文献转化为可供全文检索和分析的文本数据，这在规模上是革命性的。

二是跨语言文献的无缝解析。跨语言文献解析中的语言障碍长期制约着中国学者的西方史学史及世界史研究，这是一个无法回避的问题。而AI强大的机器翻译与多语言处理能力正以前所未有的力度打破这一瓶颈，为我们平等、高效地接入全球范围的历史文献资源提供了可能，极大地丰富了我们研究的输入。

三是多模态史料与文本的对齐。多模态大模型的能力快速发展，已经能够对图像、视频、音频等史料进行准确的解读，并能够将这些材料与文本进行语义对齐。以我们构建的古地图数字化平台^②为例，AI技术不仅能帮助我们识别、定位和检索古地图上的海量地名信息，更能通过LLM实现史料文本与古地图的多模态对齐。这意味着，我们可以将文献中一段关于路线的模糊描述直接转化为地图上一条可视的、有地理坐标的路径。这为图文互证的史学研究提供了高质量的数字化实现路径。

2. 协同优化中间层

更重要的是，AI能够协同优化中间层的运行。具备思维链（Chain of Thought, CoT）^③能力的LLMs（如DeepSeek-R1、GPT-4o等）不再仅仅是信息检索工具，更有潜力成为历史学家的思考伙伴。

一是AI辅助关联发现与逻辑推演。我们在对西夏文字典《文海》^④和OPTED英文词典的网络分析研究中，首先通过计算方法发现了词语间由互训关系构成的“强连接部件”（如OPTED中与俄罗斯宗教相关的一组生僻词）^⑤所形成的语义簇。这些是抽象的结构，理解起来仍然有难度。随后，我们要求LLM围绕这些词语讲述故事、解释关联，它成功地将抽象的数据关联，转化为富有历史情境和逻辑的叙事，帮助我们理解了结构背后的意义。这正是AI辅助中间层进行意义建构的生动体现。

二是AI驱动文献考证。在我们基于《盛宣怀档案》构建智能史料解析平台，以及利用AI考证古代地名、重构古代路网^⑥等研究中，AI的思维链能力能够帮助我们从海量碎片化信息中梳理线索、

^① Guangwei Zhang, "Knowledge Protocol Engineering: A New Paradigm for AI in Domain-Specific Knowledge Work," arXiv preprint, arXiv: 2507.02760, 2025.

^② 潘威等：《古旧地图的信息化》，《图书馆论坛》2021年第11期；潘威等：《历史地理信息化与图情研究融合的必要性与可行性——以“数字历史黄河”为中心的考察》，《图书情报知识》2021年第3期。

^③ Jason Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, Red Hook: Curran Associates Inc., 2022, Article 1800.

^④ 张光伟：《西夏文字典〈文海〉的网络分析》，《数字人文》2022年第1期。

^⑤ 张光伟：《当AI读懂我的西夏文研究》，https://m.thepaper.cn/newsDetail_forward_30923508；Zhang Guangwei, "How AI is Becoming the Perfect Partner for Historians," <https://www.sixthtone.com/news/1017267>；Guangwei Zhang, "Road Network Reconstruction Based on Place Name Extraction in Classical Chinese Texts with LLMs," *Digital Humanities* 2024, Washington: George Mason University, 2024.

^⑥ Guangwei Zhang, "Road Network Reconstruction Based on Place Name Extraction in Classical Chinese Texts with LLMs," *Digital Humanities* 2024, Washington: George Mason University, 2024；Yaoyao Zhang, et al., "Range Contraction of the Yangtze Finless Porpoise Inferred from Classic Chinese Poems," *Current Biology*, Vol. 35, No. 9, May 2025, pp. R329-R330.

进行多步推理、比对多源证据，极大地增强了中间层进行深度加工和批判性分析的能力。

3. 创新输出层的形态与出口

AI 也为输出层带来了革命。它不仅可以通过数据可视化、交互式叙事等方式创新历史研究成果的呈现形态，更能拓宽历史智慧的出口。

一是研究成果的智能呈现。在我们的研究中，AI 已经参与到基于史料的唐代交通路线考证中，并能够直接在数字化的古地图平台上进行可视化呈现，甚至可以将《李娃传》这类文学作品的情节，自动绘制到我们与 AI 协作重构的唐长安城数字地图上。这种动态的、多模态的、叙事性的输出，比传统的文字或静态地图更具表现力和传播力。

二是历史智慧的跨界赋能。中国科学院学者将古典诗词中记录的江豚信息，应用于生态保护研究的案例^①，展示出史料蕴含价值的现代体现。一个由 AI 赋能的、拥有强大元认知中间层的历史学，其跨学科知识升华能力更强，因此更能有效地以人文为源、以科技为器，在气候变化、公共治理、文化传承等复杂议题上贡献其独特的、经过深度加工的智慧。

五、历史学家的角色重塑

在 AI 的深度介入下，历史学家的核心价值与角色正经历一场深刻的重塑。正如《美国历史评论》的专题所讨论的，那种以记忆大量史实、独立完成所有文献处理为核心能力的传统学者形象正面临挑战。未来，能够在 AI 时代真正发挥价值升华作用、驾驭元认知中间层的历史学者，其核心竞争力将不再是作为知识的存储器，而是作为智慧的中央处理器（CPU），这更多地体现在以下四个相互关联的核心能力上：

1. 提出深刻问题的能力

AI 善于解答，但高质量、有穿透力的问题仍源于人类的好奇心与深刻洞察。在我们的模型中，历史学家的首要角色是为整个神经网络定义其需要解决的任务函数。AI 可以高效地处理海量输入，但输入什么、为了回答哪个根本性的问题——最初的、充满创造性的提问，是一切价值创造过程的源头。

2. 主导人机协同的能力

这种能力要求学者从技术的被动使用者转变为 AI 行为的主动设计者。这正是我们提出的知识协议工程（KPE）^② 的核心理念。通过设计和撰写知识协议（Knowledge Protocol, KP），学者实际上是在为 AI 的元认知中间层编写其核心的激活函数，即我们所论述的赋予意义之网、提供时间智慧等方法论。熟练地设计研究路径（即构建 KP），善于向 AI 提问（即调用 KP），并能批判性地评估、引导和修正 AI 的输出，这不仅是 AHA《历史教育中的人工智能指导原则》所强调的新时代核心素养^③，更是历史学家作为 AI 导师和知识架构师的价值所在。

3. 进行价值判断与伦理审思的能力

在复杂的历史情境中进行价值判断，在技术应用中保持伦理自觉，这是 AI 在当前乃至可预见的未来都无法替代的人类智慧。在我们的模型中，这对应于对中间层权重和激活逻辑的最终校准。AI 可以根据协议分析出某个政策在经济上的效率，但只有历史学家能注入人文关怀这一激活函数，去权

^① Yaoyao Zhang, et al., "Range Contraction of the Yangtze Finless Porpoise Inferred from Classic Chinese Poems," *Current Biology*, Vol. 35, No. 9, May 2025, pp. R329-R330.

^② Guangwei Zhang, "Knowledge Protocol Engineering: A New Paradigm for AI in Domain-Specific Knowledge Work," arXiv preprint, arXiv: 2507.02760, 2025.

^③ American Historical Association, "Guiding Principles for Artificial Intelligence in History Education," <https://www.historians.org/resource/guiding-principles-for-artificial-intelligence-in-history-education/>.

衡这一效率背后对不同社会群体的公平与正义意味着什么。这种能力确保了我们的神经网络输出的不仅是技术理性的结果，更是符合人类核心价值观的洞见。

4. 进行最终的意义建构与理论创新的能力

AI 可以辅助我们获得海量的信息和局部的洞见，但将这一切整合进一个更宏大、更深刻、更具人文关怀的解释框架与叙事之中，完成最终的意义建构，依然是历史学家的终极任务。这对应于对输出层的最终诠释。更重要的是，历史学家需要发挥反思性批判这一最高阶的激活函数，去审视和挑战 AI 辅助下产生的整个知识体系，从而提出全新的理论范式。这正是从利用 AI 进行研究到对 AI 时代本身进行历史学研究的飞跃。

这些都对我们培养新一代的历史学者提出了新的更高的要求：他们不仅需要深厚的史学根基，还需要开放的跨学科视野、积极的技术拥抱心态和强烈的社会责任感。

结 语

AI 为史学研究带来的既是挑战，更是历史性的机遇。正如相关学者所呼吁的那样，它迫使我们去深刻反思历史学的核心价值。我们提出的 KPE 和元认知中间层模型，正是为这种价值的显式化和可计算化提供的一种理论尝试。历史学的未来不在于抗拒 AI，也不在于被动地使用 AI，而在于主动地、智慧地与 AI 构建一种双向赋能的共生关系。在当前这个可能被后世历史学家称为“第一个真正的 AI 时代”的开端，历史学这门古老的学科将通过不断修炼和强化作为元认知中间层的独特能力，在新的知识生产体系中扮演更核心、更能动的角色，为理解我们复杂的过去、应对我们严峻的现在、启迪我们不确定的未来，贡献无可替代的、充满智慧的洞察。

责任编辑：刘 莉