

人工智能驱动的科学学理论与方法重塑

李赫扬,方思清,陈凯华

(中国科学院大学公共政策与管理学院,北京 100190)

摘要:人工智能的迅猛发展正深刻改变科学研究的运行逻辑与组织模式,也将在根本上重塑科学学的理论基础与方法体系。本文基于人工智能驱动科学研究的兴起,系统分析了人工智能赋能下科学学研究在研究对象、核心问题与方法工具上的整体转型。研究表明,人工智能既作为科学系统的变革力量,催生新的科学学研究对象与问题,又作为强大的分析工具突破传统科学学方法的技术瓶颈。这一系列变革将使科学学实现从外部表征到内在机制、从现象描述到机制解释、从被动观察到主动设计的根本性拓展,新的研究范式呈现出前瞻性、系统性、协同性与设计性等核心特征。而在方法体系层面,人工智能推动了科学学观察测量的客观化与多模态化、分析建模的预测化与网络化、解释预测的系统化与可解释化,显著增强了科学学在科研评价、知识发现与科研生态设计中的实践能力。然而,技术、伦理、制度等方面的风险与挑战也亟待应对。本文提出,应在保持科学学学科独特性的同时,以方法创新驱动范式演进,构建开放、透明、可解释且具有前瞻指导力的方法体系,以实现人工智能赋能下科学学的持续跃升。

关键词:人工智能;科学学;AI for Science;方法创新;范式变革

中图分类号:G301

文献标识码:A

人工智能正在重塑科学研究的基本范式。从蛋白质结构预测到智能化学家,人工智能驱动科学研究(AI for Science)已成为当代科学发展的重要趋势^[1,2]。科学学(Science of Science)作为科学研究体系内的一门具体学科,同样受到人工智能驱动科学研究带来的深刻影响。然而,与其他具体学科不同,科学学以整体科学研究为研究对象^[3],其自身的独特性使其受到的影响具有复合性。人工智能以方法与工具的形式,直接影响科学学研究的技术能力,也通过重塑科学研究本身的逻辑与认知方式,间接改变科学学的研究对象与问题范式。

传统科学学的研究范式长期存在两条并行路径:以科学计量学为代表的定量分析,主要依赖引文、合作网络等外部表征数据^[4];以科学哲学、科学社会学等为代表的定性思辨,侧重于案例、观察和理论构建。这些方法在揭示科学活动的某些规律上功不可没,但其内在局限性也日益凸显。例如科学计量学更多关注科学知识生产的外部轨迹,例如论文

的发表量、引用量等,却难以直接反映论证质量、思想原创性与方法贡献,且在实际应用中存在被制度性利用和异化的风险。人工智能特别是生成式人工智能和大语言模型等的崛起,为科学学提供了破局的契机。人工智能技术开始具备理解、生成和推理的能力,可直接分析科学文献的深层语义^[5],从复杂数据中学习规律并进行因果推断^[6]。这意味着,以往因技术限制而无法分析的对象(如论证逻辑、思想新颖性等)现在可以被纳入分析视野。同时,人工智能对科研实践自身的改造又产生了一系列新现象和理论问题。例如,人工智能如何改变科学发现的路径与动力结构?人机协作如何影响科研网络的形成与演化?科研活动的数据化与自动化又如何改变科学发展的可预测性与复杂性等等。

因此,人工智能对科学学的影响并非局限于技术层面的赋能,而是一次方法论与认识论的双重转变。第一,人工智能作为方法与工具,在传统的科学学研究范式中引入了更为先进的技术能力,极大拓

收稿日期:2025-08-15;修回日期:2025-12-01

基金项目:国家杰出青年科学基金项目(72025403);中国科学院战略研究与决策支持系统建设专项(GHJ-ZLZX-2025-11)

作者简介:李赫扬(1996-),男,助理研究员。方思清(2002-),女,硕士研究生。陈凯华(1980-),男,教授,通讯作者,E-mail:chen-kaihua@ucas.ac.cn。

展了数据获取、知识分析和模式识别的深度与广度;第二,人工智能作为科学系统的变革力量,正在推动整个科学研究的运行逻辑、组织结构与知识生产模式发生深刻变化,而这种上游科学体系的重构,又必然反作用于科学学的研究对象、问题意识与理论框架。人工智能与科学学的相互赋能,使得科学学有望摆脱传统定性与定量分析的局限,真正成为能够有效指导科研评价、优化资源配置、加速知识创新的强大工具。

基于上述认识,本文将从以下三个层面系统阐述人工智能驱动下科学学的重塑过程:第一,阐述人工智能如何在宏观层面重塑科学研究,为科学学转型提供背景;第二,探讨科学学在研究对象、核心问题及理论上的结构性变化,揭示新研究范式的核心特征;第三,聚焦方法体系这一关键维度,深入剖析人工智能如何在观察测量、分析建模、解释预测三个层次推动科学学方法论的革命性升级。

1 人工智能驱动的科学学研究范式转型

人工智能的浪潮正以前所未有的深度与广度席卷全球,科学研究这一人类最核心的认知活动也正经历着一场由其驱动的深刻变革。这场变革不仅体现在工具与效率的提升,更触及了科学研究的底层逻辑与范式结构。为了更系统地理解人工智能驱动的科学学研究范式转型,本章首先分析人工智能如何在科学研究体系内部形成新的认知与运行机制,从根本上驱动科学学研究范式转型;随后探讨科学学在研究对象、核心问题上的结构性变化,并总结人工智能驱动下科学学新范式的主要特征。

1.1 人工智能正在重塑科学研究的底层逻辑

要理解人工智能如何重塑科学学研究,必须首先认识到人工智能正在对整个科学研究体系进行根本性重塑。作为科学学的研究对象,科学研究本身正经历一场触及底层逻辑的深刻变革,也由此为科学学研究的范式转型提供了最根本的驱动力。人工智能驱动的科学学研究已不仅是一个技术概念,而是代表着一个全新的、由人工智能驱动的科学学研究范式。它是一个涵盖了科研目标、学术理念、协作方式乃至认知模型的集合体,其核心在于将人工智能作为一种通用赋能技术,深度融入科学发现的全过程,从而引发科学研究的结构性变革。要理解这场变革对科学学的深远意义,需要从科学研究本身正在发

生的系统性改变入手。这种改变可以从三个层次来理解。

首先是认知主体的演化。在传统科研范式中,人类科学家是唯一的认知主体。而人工智能的介入正在改写这一基本设定。人工智能在科研活动中的角色正在经历从工具到智能体(Agent)再到协作者的深刻演化。在未来的人机协同阶段,人工智能与人类智慧将实现有机融合,人工智能可能成为部分科研活动的主导者,与人类科学家共同组成一个前所未有的高效知识创造共同体^[7]。

其次是研究过程的重构。人工智能对科研活动的影响不再局限于某个环节,而是渗透到从基础研究的假设生成、实验设计,到技术开发的数据分析、模型验证,再到最终应用转化的整个创新链条中。在每一个关键节点,人工智能都发挥着放大和加速作用。它能显著压缩研发周期,降低试错成本,延展人类研究者的智力边界,从而系统性地提升整个科研生态系统的创新效能。这种全链条的深度介入,使得科学研究的内在运行逻辑发生了根本改变,科学知识的生产方式正在被重新定义。

第三是知识生产结构的变革。人工智能正在各个领域催生人工智能与学科深度融合的新研究范畴,从根本上改变着学科边界和研究方法。这种融合具有显著的跨学科效应。它不是简单的工具应用,而是促使传统学科在认识论与方法论上进行根本性迭代的催化剂。学科之间的壁垒被打破,知识生产从线性的、封闭的模式转向网络化的、开放的模式,科学研究的组织形态正在被重新定义。

这三个层次的系统性重塑,构成了科学学研究范式转型的根本驱动力。首先,当人工智能成为科研活动的有机组成部分时,科学学的研究对象必然从人类科学家的活动扩展为人机协同的知识创造过程,从单学科内的知识生产扩展到跨学科融合的创新生态。与此同时,人工智能也为科学学提供了全新的研究工具,使得对科学文本深层语义的大规模理解、对海量数据中隐藏规律的系统挖掘成为可能。正是在这种双重驱动下,科学学的研究问题得以从现象描述走向机制解释、从历史总结走向趋势预测、从被动观察走向主动设计,从而迎来了系统性的重构契机。那么,这种重构具体表现为哪些新的研究对象和新的研究问题,则需要进一步深入讨论。

1.2 科学学研究的新对象与新问题

人工智能驱动科学研究的兴起推动科学学研究

形成了全新的范式,这一范式的核心特征集中体现在研究对象和研究问题的根本性变革上。传统科学研究的对象呈现分化:以科学计量学为代表的定量研究主要集中在外部可观测表征,如论文数量、被引次数等指标^[8];以科学社会学、科学哲学为代表的定性研究则聚焦于科学共同体、实验室观察等案例。这两种路径都有一定的局限性。如科学计量学长期依赖论文数量、被引次数、期刊影响因子、合作网络规模等量化指标来描绘科学发展图景。这些外部轨迹虽然能够反映科学知识生产的数量特征和传播轨迹,但其根本局限在于无法触及科学贡献的内在价值和真实质量,研究者只能通过这些间接指标来推测学术影响。科学社会学等则较难归纳出普适规律,且分析过程依赖研究者的主观判断。人工智能技术的突破将使科学学研究对象发生历史性拓展,实现从表征层到内核层的跨越。生成式人工智能特别是大语言模型具备了理解和分析科学文本深层语义的能力,这使得过去因技术限制而无法纳入研究视野的核心对象现在可以被直接分析。论文的论证逻辑、思想的新颖程度、理论的原创性、方法的创新性、结论的可靠性等原本需要专家主观判断的内容,现将可以通过人工智能技术进行大规模、标准化的客观分析。更重要的是,人工智能驱动的科学研究模式本身也成为科学学的新研究对象。人机协同创新模式、AI Agent 参与的知识发现过程、算法化的假设生成与验证、自动化的实验设计与执行等新兴科研现象正在重塑科学活动的组织形态和认知路径,成为科学学必须纳入视野并深入解释的新对象。这种研究对象的拓展使得科学学研究不再局限于科学活动的表象层,而是能够深入到科学知识创造的内核层,真正把握科学发展的内在逻辑。

研究对象的拓展必然带来研究问题的深化。传统科学学主要停留在现象描述和统计分析层面。研究者关注的核心问题包括某个领域发表了多少论文、科研合作网络如何分布、科学家群体中的资源获取机制是什么、政府如何制定科技政策来引导和监管等等。这些问题虽然能够提供科学活动的基本图景,但本质上是对已发生现象的记录和统计以及提出建议,缺乏对深层机制的探索,难以回答科学发展的内在驱动力、演化规律、预测模式等更具洞察力和前瞻性的关键问题。人工智能技术的引入使得科学学研究能够提出和解决更具深度和复杂性的问题,实现了从现象描述到机制解释、从历史总结到趋势

预测、从被动观察到主动设计的多重跃升。人工智能技术具备强大的模式识别和规律发现能力,能够从海量复杂的科学数据中挖掘出隐藏的机制和规律,研究者将可以探讨过去因技术限制而难以触及的深层问题。这种问题导向的根本转变使得科学学研究从简单的现象统计转向了机制理解,从静态的截面分析转向了动态的过程探索,从经验性的归纳总结转向了理论性的因果推断,从被动的观察记录转向了主动的设计干预。

科学学研究对象和问题的拓展,正是人工智能在工具层面和系统层面共同作用的结果。在工具层面,生成式人工智能如大语言模型为科学学提供了直接分析科学文本语义、论证逻辑、创新程度的技术可能性;在系统层面,人工智能驱动的科学研究模式变革产生了大量新的研究现象和问题,这些新对象、新问题反过来又推动科学学不断拓展理论边界,发展新的分析框架。两个层面的互动形成正反馈循环:新工具使新对象可研究,新对象催生新问题,新问题驱动新理论,新理论又指导新工具的开发与应用,使得科学学的研究边界得以前所未有地拓展。

1.3 人工智能驱动下科学学研究范式的核心特征

人工智能通过拓展研究对象与深化研究问题,为科学学研究开辟了全新的空间。然而,对象与问题的变革最终要落实到认识论与方法论的整体转型上,才能真正形成新的研究范式。理论范式的转换是学科发展的重要标志。人工智能不仅改变科学学的研究对象与问题范畴,更将在根本上重塑“如何认识科学”和“如何研究科学”的整体逻辑。这种新范式的内核,不在于单一技术手段的替换,而在于科学学在认知能力、分析视野、研究主体和实践功能等多个维度的整体性重构,从而使其成为一个更具前瞻性、系统性、协同性与设计性的知识体系。

第一,新范式具有显著的前瞻性。区别于传统范式的事后解释,新范式将具备了事前预测的强大能力。通过对海量科学数据的深度挖掘和模式识别,人工智能可识别科学发展的早期信号,预测新兴研究领域的出现,预判学科交叉融合的趋势,预估科研投入的产出效果等等。这种预测能力的获得,使得科学学研究从传统的“历史学科”转向了具有前瞻性的“未来学科”,为科学规划和科研决策提供了重要的智能支持^[9]。第二,新范式凸显系统性。传统科学学研究往往是碎片化的:科学计量学关注论文和引用,科学社会学关注人际网络,科学政策研究

关注制度设计等等,各个分支领域相对独立,难以形成对科学系统的整体认识。这种碎片化的根本原因在于数据获取和处理能力的限制,研究者只能聚焦于某一个侧面进行深入分析。人工智能技术的引入,使科学学能够在统一的分析框架中整合多尺度、多类型、多来源的数据,实现从微观个体(科研人员、项目)到中观组织(团队、机构、学科)再到宏观体系(国家、全球科研网络)的全景化观测^[10]。这种系统性不仅体现在研究对象的覆盖面,更体现在分析逻辑的互联性——能够揭示不同层次之间的反馈机制和演化耦合,从而更准确地刻画科学发展的整体动力学。第三,新范式带来了协同性。传统科学学研究主要依赖人类专家的经验判断和理论思辨,而新范式则建立在人机深度协同的基础之上^[11]。在这种协同模式中,人工智能系统承担数据处理、模式识别、初步分析等计算密集型任务,人类研究者则发挥创造性思维、价值判断、理论建构等不可替代的作用。同时,人类与人工智能也能协作完成论文同行评议等工作^[12]。这些协同不是简单的分工,而是人类智慧与机器智能的有机融合:人工智能的强大计算能力弥补了人类在处理海量数据方面的局限,而人类的洞察力、直觉和创造力则为人工智能分析提供了深层次的理解和价值导向^[13]。人机协作的新模式重塑了科学学研究的学术生态^[14]。最后,新范式具有更强的设计性。传统科学学往往局限于“理解科学”,而在人工智能赋能下,科学学不仅能够识别问题,还能够为科研制度优化、科研组织架构设计、合作网络构建、人才培养路径规划等提供可操作的方案。这种从观察者到参与者的角色转变,意味着科学学在未来将不仅是科学发展的镜子,更是塑造科学生态的重要力量。

综上,人工智能驱动的科学学研究新范式,不仅在研究对象和问题导向上发生了根本转变,更在认识论和价值取向上完成了深刻升级,使科学学从被动的记录者与解释者,成长为能够预测未来、整合系统、协同智能、主动设计的科学发展参与者。然而,这些范式特征的实现,根本上依赖于方法体系的重塑。预测性需要从描述统计到预测建模的跨越,系统性依托于从单一维度到多模态融合的观测能力,协同性建立在人机协作的新型分析框架之上,而设计性则源于从相关分析到因果推断的机制洞察。可以说,方法创新是连接理论愿景与实践能力的关键桥梁:没有观察测量的智能化,就无法获取支撑新范

式的高质量数据;没有分析建模的计算化,就难以从复杂现象中提取深层规律;没有解释预测的智能化飞跃,就无法实现从理解过去到设计未来的跨越。

因此,本文将聚焦方法体系这一核心维度,沿着科学认识深化的逻辑路径,从数据驱动的描述层到模型驱动规律层的规律层,再到理论驱动的洞察层,系统剖析人工智能如何在观察测量、分析建模、解释预测三个层次重塑科学学的技术基础与实践路径,从而为前文所述的范式转型提供坚实的方法论支撑。

2 科学学方法体系的智能化重塑

如前所述,人工智能驱动的科学学新范式在预测性、系统性、协同性和设计性等方面展现出显著特征,而这些特征的实现根本上依赖于方法体系的重塑。人工智能与科学学的融合不仅仅是技术工具的简单叠加,而是一场深层次的方法论革命,它正在重新定义我们理解科学活动本质、探索知识演化规律、评估科研影响的基本框架。这种变革并非孤立地发生在某一技术层面,而是沿着科学认识深化的逻辑路径展开,实现了从观察测量到分析建模、再到解释预测的递进跃迁,如图1所示。

这一演化路径揭示了科学学方法体系在人工智能赋能下的系统性重构:人工智能推动观察与测量走向客观化,推动分析与建模实现计算化,并最终促成解释与预测的智能化飞跃。本章节将依次从这三个层次展开论述。

2.1 观察与测量方法的智能化变革

科学研究的起点是观察和测量,传统科学学主要依靠研究者的主观观察和有限的统计工具来收集和处理数据。人工智能技术的引入从根本上改变了科学学研究中数据获取、处理和分析的基本方式,实现了从主观到客观、从静态到动态、从单一到多模态的全面转型。

首先,人工智能促进了从主观观察到客观量化的演化。传统科学学研究往往依赖研究者的主观判断和经验积累。例如,在分析学科发展趋势时,研究者需要通过阅读大量文献、参与学术会议、访谈专家等方式获取信息,然后基于个人理解和经验进行主观判断。这种方法不仅效率低下,而且容易受到研究者的知识背景、认知偏见、文化立场等因素的显著影响,难以保证观察结果的客观性和可重复性。人工智能及大数据技术为科学学研究提供了全新的客

观化测量工具。大规模文本挖掘技术能够自动处理数百万篇学术论文,精确提取其中的关键信息,如研究主题、方法类型、合作模式^[15]、引用关系^[16]等,并将这些信息转化为标准化的量化指标^[17]。例如,Tshitoya 等开发的无监督词嵌入技术能够从材料科学文献中自动捕获潜在的科学知识,发现了传统方法难以识别的材料属性关联和研究趋势^[18]。知识图谱构建技术进一步推动了这一客观化进程,它能够将传统需要专家主观判断的学科知识结构转化为客观的图谱表示。通过实体识别、关系抽取、图谱融合等技术,研究者能够构建涵盖概念、方法、理论、应用

等多层次的科学知识图谱^[19]。如 CS - KG 2.0 项目构建的大规模计算机科学知识图谱就是这一技术应用的典型范例,该图谱整合了海量学术资源,为计算机科学领域的知识结构分析提供了客观化的数据基础^[20]。这种客观化测量的意义在于,它将原本高度主观、难以量化的科学活动转化为可计算、可比较、可验证的数字化信息。研究者不再需要依赖个人的有限经验和主观判断,而可以基于大规模、高精度的客观数据进行科学分析^[21]。这不仅显著提高了研究结果的可信度和可重复性,也为发现新的规律和模式提供了可能。

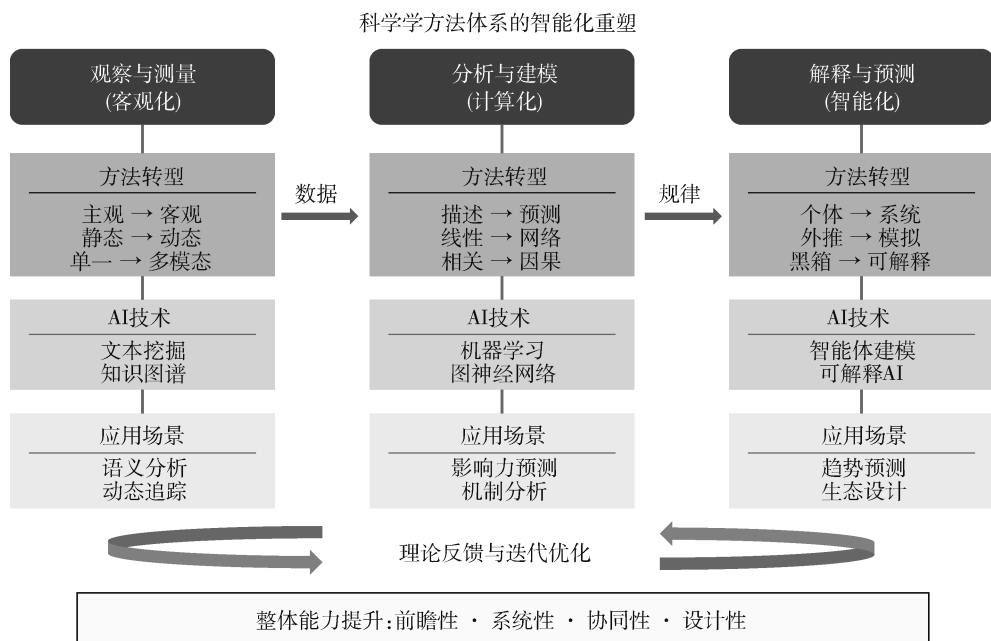


图 1 人工智能赋能科学学研究方法的重塑

Figure 1 AI-enabled reshaping of research methodologies in the Science of Science

第二,人工智能助力了从静态分析到动态追踪的突破。传统科学学研究常采用横截面分析方法,即在特定时间点对科学现象进行静态描述,难以捕捉科学活动的动态演化过程。这种方法的局限性在于无法揭示科学发展的内在机制和驱动力量,更无法预测未来的发展趋势。人工智能赋能的时间序列分析和动态网络建模技术,加速了科学学研究从静态描述到动态追踪的转变。通过构建时变知识网络,研究者能够实时追踪概念关系的演化、学科边界的变迁、研究热点的兴衰更替^[22]。深度学习模型能够从历史数据中学习科学发展的内在规律,识别出驱动变化的关键因素和影响机制^[23]。Gao 等人开发了一套系统性的测量框架来量化人工智能在科学

研究中的直接使用和相关效益^[24]。该研究通过动态追踪发现,人工智能的使用和效益在各个科学领域中呈现广泛分布,且自 2015 年以来增长尤为迅速。这种动态分析为理解人工智能技术在科学研究中的扩散机制和影响模式提供了重要洞察。

第三,人工智能实现了从单一维度到多模态数据的融合。传统科学学研究往往局限于单一类型的数据源,如仅依靠引文数据分析知识结构与期刊水平,或仅通过问卷调查了解科研人员行为。这种单一维度的分析方法难以全面反映科学活动的复杂性和多样性。人工智能技术使得多模态数据的综合分析成为可能。现代科学学研究能够同时处理文本、图像、网络、时间序列等多种类型的数据,通过多模

态融合技术将不同来源、不同格式的信息整合为统一的分析框架。例如 Liu 等人开发了融合深度学习和网络科学的计算方法,构建了艺术作品、电影和科学出版物的高维表征,通过文本、图像、网络等多模态数据分析揭示了创造性职业生涯中“热点期”(高影响力作品集中爆发期)的普遍规律^[25,26]。更重要的是,人工智能技术不仅能够处理传统的结构化数据,还能够有效处理大量非结构化数据,如学术会议的视频记录、研究者的社交媒体发言、实验室的日常活动记录等。这些软信息虽然难以直接量化,但往往包含着对理解科学活动至关重要的背景信息和隐含知识。通过自然语言处理、情感分析、行为识别等技术,人工智能系统能够从这些非结构化数据中提取有价值的洞察,为科学学研究提供更加全面和深入的信息基础。

2.2 分析与建模方法的计算化转型

在获得丰富的观察数据基础上,科学学研究的核心任务是通过分析和建模来发现数据背后的规律和机制。人工智能技术不仅提供了更强大的计算工具,更重要的是引入了全新的分析范式,实现了从描述统计到预测建模、从线性关系到复杂网络、从相关分析到因果推断的根本性转型。

首先是从描述统计到预测建模的分析升级。传统科学学研究主要采用描述性统计方法及回归模型来描述科学现象的基本特征。这种方法虽然能够提供现象的基本描述,但缺乏预测能力和解释深度,难以全面回答“为什么”和“接下来会怎样”等更深层次的问题。机器学习技术的引入使得科学学研究能够构建复杂的预测模型,与传统统计方法相比具有更强的非线性建模能力和更好的泛化性能。例如,深度神经网络能够自动发现数据中的复杂模式^[27],随机森林算法能够处理高维特征空间中的复杂交互效应,支持向量机能够在非线性空间中寻找最优分类边界。这种从描述到预测的转变不仅提高了分析的技术水平,更重要的是使科学学研究不再满足于简单描述科学现象,而是力图构建能够预测和解释的理论模型。例如,Krenn 等人从约 75 万篇科研论文构建语义网络,并结合深度神经网络预测科学概念间未来连接和研究趋势,并识别结构洞中的潜在学科交叉^[28]。结果表明该方法能高精度预测未来五年内的科研发展趋势,并生成个性化的研究假说。

第二是从线性关系到复杂系统的认知改变。传统科学学研究往往基于线性思维,假设变量之间存

在简单的因果关系或相关关系。这种简化的分析框架虽然便于理解和操作,但难以捕捉科学系统的复杂性和非线性特征。现实中的科学活动往往涉及多个行为主体、多层次结构、多重反馈机制,呈现出典型的复杂系统特征。网络科学和复杂系统理论的引入为科学学研究提供了全新的分析框架^[29]。研究者开始将科学系统建模为复杂网络,其中节点代表不同的科学实体(如研究者、机构、概念、论文等),边代表它们之间的各种关系(如合作、引用、影响等)。通过分析网络的拓扑结构、动态演化、传播机制等特征,研究者能够深入理解科学系统的组织原理和运行规律。图神经网络等先进技术进一步提升了网络分析的能力。与传统网络分析方法相比,图神经网络能够自动学习节点和边的表示,捕捉网络中的高阶模式和隐含结构。

第三是从相关分析到因果推断的机制探索。科学研究的目標之一是理解现象背后的因果机制,而非简单的相关关系。然而,传统科学学研究往往局限于相关分析,难以建立可靠的因果关系。这主要是因为观察性数据中存在大量的混杂变量、选择偏差、反向因果等问题,使得因果推断变得较为困难。人工智能驱动的因果推断技术将为这一问题提供了创新解决方案。例如已有研究利用文本嵌入和匹配方法,通过因果推断的视角衡量论文的科学影响力^[30]。因果发现算法能够从观察数据中自动学习变量之间的因果结构,识别出真正的因果关系和虚假的相关关系。反事实推理技术能够估计在不同干预条件下可能产生的结果,为政策评估和决策支持提供科学依据。

2.3 解释与预测方法的智能化飞跃

如果说观察、测量、分析和建模是理解科学系统的基础,那么解释和预测则是科学学研究价值的最终体现。人工智能技术正在推动科学学研究实现从个体预测到系统性发现、从线性外推到非线性模拟、从“黑箱”预测到“白箱”解释的智能化飞跃,从而更深刻地揭示科学发展的内在逻辑,并前瞻性地把握未来科技脉搏。

首先是从个体预测到系统性发现的视野拓展。传统科学学研究的预测目标往往聚焦于孤立的实体,如预测某位科学家的学术潜力^[31]、某项技术的成功概率。然而,对于整个科学体系而言,更具战略价值的是发现那些尚未被充分挖掘、潜力巨大的“知识无人区”和“技术蓝海”。人工智能技术,特别

是基于知识图谱和大规模语言模型的分析方法,能够将积累的全部科学文献、专利和数据集构建成一个宏大的、动态演化的知识空间。通过在这个空间中进行拓扑分析和语义挖掘,使得人工智能从一个研究科学的分析工具,转变为一个启发科学的灵感引擎。例如,Peng 等人在超过 5300 万篇论文的引用网络上进行随机游走,将游走路径视为句子并输入改良的 word2vec 模型,生成连续的期刊向量以构建高分辨率的学科组织图谱。研究发现,这种神经嵌入方法相比传统稀疏模型更能准确捕捉期刊间相似性及跨学科联系^[32]。Arts 等利用自然语言处理提取专利中的词组组合构建文本嵌入的新指标,用于衡量专利的新颖性与影响力,在识别技术突破方面表现明显优于传统的引文和分类方法^[33]。

第二是从线性外推到非线性演化模拟的突破。传统的科学发展预测方法多依赖于趋势外推,即假设未来的发展是过去趋势的线性延伸。这种方法在稳定发展的时期或许有效,但无法捕捉科学发展中固有的非线性、突变和涌现现象,如范式革命、颠覆性技术的诞生以及知识的爆炸性增长。人工智能驱动的复杂系统建模,特别是基于智能体建模,为科学学研究提供了全新的模拟和预测范式。研究者可以构建一个虚拟的科学世界,其中的智能体代表着科学家、科研机构、资助方等行为主体。这些智能体被赋予了基于真实世界数据学习而来的行为规则和决策模型^[34]。通过大规模模拟,研究者可以观察在不同政策干预、资源配置或合作模式下,整个科学系统将如何演化。这种方法实现了从“预测”到“预演”的转变。例如,Gao 等人探讨了大语言模型如何增强智能体建模的能力,尤其是在社会、物理和混合环境中的应用,大语言模型能够帮助模拟智能体在复杂社会网络中的决策过程,提升模拟的真实性和复杂性。这使得研究者能够更好地捕捉科学系统中的非线性交互和涌现现象^[35]。这种非线性的演化模拟能力,为科技战略和政策的制定提供了前所未有的实验平台。

第三是从“黑箱”预测到“白箱”解释的认知深化。许多传统的预测模型,包括部分早期机器学习模型,虽然能够达到较高的预测精度(例如,预测某篇论文未来将被高引),但其决策过程往往是一个不透明的“黑箱”,无法解释其做出特定预测的具体原因。这使得研究者难以将预测结果转化为可理解的知识和可信的理论。可解释性人工智能的发展有

望在很大程度上缓解这一难题。诸如 SHAP (SHapley Additive exPlanations)、LIME (Local Interpretable Model - agnostic Explanations) 等技术,能够打开机器学习模型的“黑箱”,量化分析每个输入特征(如作者声誉、期刊影响力、主题新颖性、研究的跨学科性)对最终预测结果的贡献度。这一飞跃的意义在于,它能够将科学学的分析范式从“预测什么”深化到了“解释为什么”。研究者不仅可以预测科学影响力,更能够构建和验证关于科学影响力驱动机制的理论。这种“白箱”解释能力,使人工智能驱动的科学学研究不再仅仅是数据挖掘,而是真正能够产生深刻洞见和理论贡献的科学探索。

3 结论与展望

3.1 主要结论

本研究系统揭示了人工智能驱动下科学学研究的范式转型及其实现路径。研究表明,人工智能对科学学的影响呈现双重驱动机制:一方面,人工智能作为科学系统的变革力量,正在重塑科学研究的运行逻辑与组织模式,催生出人机协同创新、算法辅助发现等新科研现象,这些成为科学学必须纳入视野的新研究对象;另一方面,人工智能作为强大的分析工具,突破了传统方法的技术瓶颈,使得对科学活动内核的直接观测、对知识演化规律的精准建模、对未来趋势的可靠预测成为可能。在这种双重驱动下,科学学研究经历了从对象到问题、从问题到范式、从范式到方法的系统性重构。

在研究内容层面,科学学实现了从外部表征到内在机制、从现象描述到机制解释、从被动观察到主动设计的根本性拓展。在范式特征层面,科学学形成了以前瞻性、系统性、协同性和设计性为核心的新研究范式。在方法体系层面,人工智能推动科学学实现了三个层次的重塑:观察与测量层面,从主观到客观、从静态到动态、从单一模态到多模态融合;分析与建模层面,从描述统计到预测建模、从线性关系到复杂网络、从相关分析到因果推断;解释与预测层面,从个体预测到系统性发现、从趋势外推到非线性模拟、从黑箱预测到可解释分析。这些方法创新是连接范式愿景与实践能力的关键桥梁,为科学学在科研评价、知识发现与科研生态设计中的应用提供了坚实支撑。

理论上,本文提出了人工智能赋能科学学的整

体框架,揭示了研究对象拓展、问题深化、范式转型与方法创新之间的内在逻辑关系,强调技术革新与理论建构的双向驱动;实践上,为智能化科研评价、前瞻性知识发现与系统性科研生态优化提供了可扩展的方法路径。

3.2 挑战与发展展望

展望未来,人工智能驱动的科学学研究正站在前所未有的历史机遇面前。技术层面,大语言模型的快速发展将使科学文献的理解和分析达到前所未有的深度,多模态人工智能技术的成熟将实现对科学活动的全方位感知,具身智能与认知智能的突破将赋予人工智能系统更强的科学推理与假设生成能力。这些技术进步不仅将拓展科学学的分析边界,更将推动研究从研究科学向参与科学、从理解规律向启发发现的深度转变。然而,这一变革过程也伴随着多重挑战,需要科学学研究者保持清醒认识并采取系统性应对策略。

在技术层面,首先,人工智能在科学学中的应用可能面临算法偏见与数据质量的挑战。科学学研究依赖大规模历史数据对科研活动进行分析与评价,若原始数据内嵌了性别、地域等系统性偏差,人工智能模型则可能无意中放大这些偏见,从而影响科研评价的公正性与资源配置的公平性。其次,人工智能模型的“黑箱”特性可能导致应用中的信任危机。许多复杂的深度学习模型,在决策过程缺乏透明度和可解释性,这在对科研进行关键性判断时,将引发学术界的信任危机,限制其应用深度。

在伦理层面,首先,人工智能的自动化假设生成和研究方向预测可能对科学学研究独立性构成威胁。研究者可能会在选择课题时出现过度依赖人工智能的情况,从而忽视了人类的创新意识与学术判断。此外,人工智能的普及可能导致科学学研究的资源分配失衡。随着人工智能工具使数据处理、基础分析等技术环节的获取成本大幅降低,原本由技术稀缺性带来的价值被稀释。这种转变使得非技术性资本(如高质量数据等)的重要性日益凸显。由此,科研体系内部可能出现基于数据与资源的新型不平等结构:拥有这些初始资源优势的顶级研究机构和高影响力学者,能够更高效地整合和利用人工智能,加速科研成果的产出和积累,从而进一步拉大学者的“马太效应”。

在制度层面,人工智能通过改变科学学研究范式,对现有科研治理体系构成挑战。一方面,它冲击

着同行评审、论文署名及成果评价等传统学术规范,其辅助甚至替代部分工作的能力,让人工智能贡献度认定、学术不端界定,以及平衡自动化评价效率与科研复杂性、创新性等问题面临新挑战。另一方面,人工智能用于项目审批、人才选拔等高危决策时,若模型输出偏差、产生偏见或引发不公,受“黑箱”特性和复杂决策链影响,难以明确模型开发者、使用者与算法本身的行政及法律责任,这种模糊性会损害科研治理的公正性与公信力,亟需新法律和行政框架界定权责边界。

面对机遇与挑战并存的未来,科学学研究需要在多重张力中寻求动态平衡:在拥抱人工智能技术的同时保持学科独特性,避免成为纯粹的技术应用领域;在追求分析精度的同时坚持理论深度,确保产出具有学术价值的理论贡献;在推动方法创新的同时强化伦理约束,使技术进步服务于科学事业的健康发展;在开放数据共享的同时维护研究自主,在国际合作中保持学术话语权。只有在技术与理论、创新与传承、效率与公平、开放与自主的有机统一,科学学研究才能在人工智能时代实现真正的跃升,成为理解科学、改进科学、引领科学的重要力量,为建设科技强国和推动科学事业发展作出独特而深远的贡献。

参考文献:

- [1] Wang H, Fu T, Du Y, et al. Scientific discovery in the age of artificial intelligence [J]. Nature, 2023, 620 (7972): 47-60.
- [2] 李国杰. 智能化科研(AI4R):第五科研范式[J]. 中国科学院院刊, 2024, 39(1): 1-9. Li G J. Intelligent scientific research (AI4R): The fifth paradigm of scientific research [J]. Bulletin of the Chinese Academy of Sciences, 2024, 39(1): 1-9.
- [3] Fortunato S, Bergstrom C T, Börner K, et al. Science of science [J]. Science, 2018, 359(6379): eaao0185.
- [4] 赵蓉英,许丽敏. 文献计量学发展演进与研究前沿的知识图谱探析[J]. 中国图书馆学报, 2010(5): 60-68. Zhao R Y, Xu L M. Bibliometrics development and evolution and knowledge mapping analysis of research frontiers [J]. Journal of Library Science in China, 2010(5): 60-68.
- [5] Zhuang Z, Chen J, Xu H, et al. Large language models for automated scholarly paper review: A survey [J]. Information Fusion, 2025, 124: 103332.
- [6] Clauset A, Larremore D B, Sinatra R. Data-driven

- predictions in the science of science [J]. *Science*, 2017, 355(6324): 477–480.
- [7] 中国科学院创新范式研究组. 加快人工智能驱动的知识创新体系建设[J]. *中国科学院院刊*, 2025, 40(10): 1709–1719. Research Group on Innovation Paradigm of Chinese Academy of Sciences. Accelerating the construction of an AI-driven knowledge innovation system [J]. *Bulletin of the Chinese Academy of Sciences*, 2025, 40(10): 1709–1719.
- [8] Liu L, Jones B F, Uzzi B, et al. Data, measurement and empirical methods in the science of science [J]. *Nature Human Behaviour*, 2023, 7(7): 1046–1058.
- [9] Krenn M, Buffoni L, Coutinho B, et al. Forecasting the future of artificial intelligence with machine learning-based link prediction in an exponentially growing knowledge network [J]. *Nature Machine Intelligence*, 2023, 5(11): 1326–1335.
- [10] Zhang Y, Zhang C, Mayr P, et al. An editorial of “AI + informetrics”: Multi-disciplinary interactions in the era of big data [J]. *Scientometrics*, 2022, 127(11): 6503–6507.
- [11] Correia A, Grover A, Jameel S, et al. A hybrid human-AI tool for scientometric analysis [J]. *Artificial Intelligence Review*, 2023, 56(S1): 983–1010.
- [12] Liang W, Zhang Y, Cao H, et al. Can Large Language Models provide useful feedback on research papers? A large-scale empirical analysis [J]. *NEJM AI*, 2024, 1(8): A10a2400196.
- [13] Wang D, Weisz J D, Muller M, et al. Human-AI collaboration in data science: Exploring data scientists’ perceptions of automated AI [J]. *Proc. ACM Hum. - Comput. Interact.*, 2019, 3(CSCW).
- [14] Yang Y, Youyou W, Uzzi B. Estimating the deep replicability of scientific findings using human and artificial intelligence [J]. *Proceedings of the National Academy of Sciences*, 2020, 117(20): 10762–10768.
- [15] Zeng A, Fan Y, Di Z, et al. Fresh teams are associated with original and multidisciplinary research [J]. *Nature Human Behaviour*, 2021, 5(10): 1314–1322.
- [16] Miura T, Asatani K, Sakata I. Large-scale analysis of delayed recognition using sleeping beauty and the prince [J]. *Applied Network Science*, 2021, 6(1).
- [17] Lin Z, Yin Y, Liu L, et al. SciSciNet: A large-scale open data lake for the science of science research [J]. *Scientific Data*, 2023, 10(1): 315.
- [18] Tshitoyan V, Dagdelen J, Weston L, et al. Unsupervised word embeddings capture latent knowledge from materials science literature [J]. *Nature*, 2019, 571(7763): 95–98.
- [19] Peng C, Xia F, Naseriparsa M, et al. Knowledge graphs: Opportunities and challenges [J]. *Artificial Intelligence Review*, 2023, 56(11): 13071–13102.
- [20] Dessì D, Osborne F, Buscaldi D, et al. CS-KG 2.0: A large-scale knowledge graph of computer science [J]. *Scientific Data*, 2025, 12(1): 964.
- [21] 郑碧丽, 侯剑华. 数智驱动的科学学: 国内理论科学研究综述 [J]. *科学学研究*, 2025, 43(3): 514. Zheng B L, Hou J H. Data-driven science of science: Review on research of domestic theoretical science of science [J]. *Studies in Science of Science*, 2025, 43(3): 514.
- [22] Sinatra R, Deville P, Szell M, et al. A century of physics [J]. *Nature Physics*, 2015, 11(10): 791–796.
- [23] Xie Y, Pan Y, Xu H, et al. Bridging AI and science: Implications from a Large-Scale Literature analysis of AI4Science [J/OL]. *arXiv*, 2024[2025-07-25]. <https://arxiv.org/abs/2412.09628>.
- [24] Gao J, Wang D. Quantifying the use and potential benefits of artificial intelligence in scientific research [J]. *Nature Human Behaviour*, 2024, 8(12): 2281–2292.
- [25] Liu L, Dehmamy N, Chown J, et al. Understanding the onset of hot streaks across artistic, cultural, and scientific careers [J]. *Nature Communications*, 2021, 12(1): 5392.
- [26] Liu L, Wang Y, Sinatra R, et al. Hot streaks in artistic, cultural, and scientific careers [J]. *Nature*, 2018, 559(7714): 396–399.
- [27] Samek W, Montavon G, Lapuschkin S, et al. Explaining deep neural networks and beyond: A review of methods and applications [J]. *Proceedings of the IEEE*, 2021, 109(3): 247–278.
- [28] Krenn M, Zeilinger A. Predicting research trends with semantic and neural networks with an application in quantum physics [J]. *Proceedings of the National Academy of Sciences*, 2020, 117(4): 1910–1916.
- [29] Zeng A, Shen Z, Zhou J, et al. The science of science: From the perspective of complex systems [J]. *Physics Reports*, 2017, 714–715: 1–73.
- [30] Kumar I, Jin Z, Mokhtarian E, et al. CausalCite: A causal formulation of paper citations [J/OL]. 2024. <https://arxiv.org/abs/2311.02790>.
- [31] Sinatra R, Wang D, Deville P, et al. Quantifying the evolution of individual scientific impact [J]. *Science*, 2016, 354(6312): aaf5239.

[32]

Peng H, Ke Q, Budak C, et al. Neural embeddings of scholarly periodicals reveal complex disciplinary organizations [J]. Science Advances, 2021, 7 (17): eabb9004.

[33]

Arts S, Hou J, Gomez J C. Natural language processing to identify the creation and impact of new technologies in patent text: Code, data, and new measures [J]. Research Policy, 2021, 50(2) : 104144.

[34]

Xi Z, Chen W, Guo X, et al. The rise and potential of Large Language Model based agents: A survey[J/OL]. 2023. <https://arxiv.org/abs/2309.07864>.

[35]

Gao C, Lan X, Li N, et al. Large Language Models empowered agent – based modeling and simulation: A survey and perspectives [J]. Humanities and Social Sciences Communications, 2024, 11(1) : 1 – 24.

Artificial Intelligence – driven theoretical and methodological transformation of the Science of Science

LI He – yang, FANG Si – qing, CHEN Kai – hua

(School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: The rapid advancement of Artificial Intelligence (AI) is profoundly changing the operational logic and organizational models of scientific research, fundamentally reshaping the theoretical foundation and methodological system of the Science of Science (SciSci). Based on the rise of AI – driven scientific research, this paper systematically analyzes the holistic transformation of SciSci research across its objects, core problems, and methodological tools under the empowerment of AI. The study indicates that AI acts simultaneously as a revolutionary force within the scientific system, generating new SciSci research objects and questions, and as a powerful analytical tool, breaking through the technical bottlenecks of traditional SciSci methodologies. This series of transformations will enable SciSci to achieve a fundamental expansion: from external representation to intrinsic mechanisms, from phenomenological description to mechanistic explanation, and from passive observation to active design. The new research paradigm exhibits core characteristics such as prospectiveness, systematicness, synergy, and designability. At the methodological level, AI promotes the objectification and multi – modality of SciSci’s observation and measurement, the predictive and networked nature of its analysis and modeling, and the systemic and explainable nature of its explanation and prediction. This significantly enhances SciSci’s practical capabilities in research evaluation, knowledge discovery, and scientific ecosystem design. Nevertheless, risks and challenges in technical, ethical, and institutional aspects urgently need to be addressed. This paper proposes that, while maintaining the unique identity of the SciSci discipline, we should drive paradigm evolution through methodological innovation, and construct an open, transparent, explainable, and prospectively guiding methodological system to achieve the continuous leap forward of SciSci empowered by AI.

Key words: Artificial Intelligence; science of science; AI for Science; methodological innovation; paradigm transformation