

# 人工智能赋能人文社会科学研究方法变革

苏新宁<sup>1</sup> 吕先竞<sup>2</sup>

1. 南京大学信息管理学院 江苏南京 210023; 2. 四川省社科重点研究基地西华大学四川学术成果分析与应用研究中心  
四川成都 610039

**摘要:**以 AlphaFold 和 ChatGPT 为代表的人工智能,正在改变整个人类社会,驱动科学研究也驱动人文社会科学研究新范式的形成。“人工智能驱动的人文社会科学研究(AI4SS)”又称人文社科研究第五范式。与分别基于实验、基于理论、基于仿真、基于数据驱动的科学第一、二、三、四范式不同,人文社科研究第五范式具有 AI 全学科全过程赋能人文社会科学研究、机器涌现的智能成为人文社会科学研究的组成部分、利用 AI 概率统计模型解决社会科学研究中的组合爆炸问题、AI 为人文社会科学研究构筑多节点的复杂网络、AI 虚拟现实促进人文社会科学理论与实践结合以及以大模型为特征的科研大平台成为人文社会科学研究的必需等六大特点。AI4SS 引发人文社会科学领域的革命,在人文社会科学研究中全面驱动选题与学术史梳理、假设生成及假设论证的方法变革。为加快推进 AI 赋能人文社会科学研究方法变革,加快面向其研究的开源开放大模型建设是必需的。

**关键词:**人工智能;人文社会科学研究;AI4SS;研究范式;研究方法变革;学术史梳理;假设生成;假设论证;开源基础设施

中图分类号:C1; TP18 文献标志码:A 文章编号:1672-8505(2025)01-0001-10

doi: 10.12189/j.issn.1672-8505.2025.01.001

## The Transformation of Research Methods in the Humanities and Social Sciences Empowered by Artificial Intelligence

SU Xin-ning<sup>1</sup> LYU Xian-jing<sup>2</sup>

1. School of Information Management, Nanjing University, Nanjing, Jiangsu, 210023, China;  
2. Sichuan Academic Achievement Analysis and Application Research Center, Xihua University, Chengdu, Sichuan, 610039, China

**Abstract:** Artificial intelligence, exemplified by tools like AlphaFold and ChatGPT, is transforming society and driving the emergence of new research paradigms, including in the humanities and social sciences. "AI-driven Research in the Humanities and Social Sciences" (AI4SS) is recognized as the fifth paradigm in these fields. Unlike the previous four paradigms of scientific research—based on experimentation, theory, simulation, and data-driven approaches, respectively—the fifth paradigm in the humanities and social sciences is defined by six key features: (1) AI empowers the entire research process across all disciplines; (2) machine-generated intelligence is becoming an integral component of humanities and social science research; (3) AI-powered probabilistic models help address

收稿日期: 2024-11-05

基金项目: 国家社会科学基金重大项目“面向国家战略的情报学教育与发展研究”(项目编号: 20&ZD32)。

作者简介: 苏新宁, 男, 教授, 博士生导师, 主要从事智能信息处理与科学评价研究。

吕先竞, 男, 教授, 四川省社科重点研究基地西华大学四川学术成果分析与应用研究中心研究员, 主要从事信息资源管理研究。

引用格式: 苏新宁, 吕先竞. 人工智能赋能人文社会科学研究方法变革[J]. 西华大学学报(哲学社会科学版), 2025, 44(1): 1-10.

the problem of combinatorial explosion in social science research; (4) AI constructs complex, multi-node networks to advance research in these fields; (5) AI-driven virtual reality fosters the integration of theory and practice in humanities and social sciences; and (6) large-scale, model-based research platforms have become essential for humanities and social science research. AI4SS is sparking a revolution, fundamentally reshaping research methodologies related to topic selection, academic history review, hypothesis generation, and hypothesis testing. To accelerate this transformation, it is critical to advance the development of open-source, large-scale models specifically designed for these disciplines.

**Key words:** artificial intelligence; research in the humanities and social sciences; AI4SS; research paradigms; transformation of research methods; academic history review; hypothesis generation; hypothesis testing; open source infrastructure

当前,人工智能正在改变整个人类社会,丰富人们的文化生活,促进科技创新与发展,是各行各业新质化的促进器。对人文社会科学领域而言,由人工智能驱动的人文社会科学研究(AI4SS)正在成为其研究的第五范式。AI4SS 引发人文社会科学领域的革命,驱动人文社会科学研究方法变革,提高研究人员选题与学术史梳理、假设生成及假设论证的科学性及效率。

本文首先对人工智能的发展演化轨迹进行梳理,接着探寻“人工智能驱动人文社会科学研究(AI4SS)”的特点,在此基础上,再论述人工智能如何赋能人文社会科学研究选题、学术史梳理、假设生成及假设论证。为促进 AI 赋能研究方法变革,本文在最后,探讨了为人文社会科学研究构建开源基础设施的必要性。

本文旨在通过尽可能全面地展示人工智能赋能人文社会科学研究方法变革的成果,进一步推动人工智能全学科、全过程、全方位应用于人文社会科学研究,有力促进人文社会科学创新发展。

## 一、人工智能的发展演化

1950 年,阿兰·图灵(Alan Turing)提出图灵测试,该测试的核心思想是通过判断一个机器能否展现出与人类不可区分的智能行为来定义人工智能。1956 年,约翰·麦卡锡(John McCarthy)在达特茅斯会议上首次提出人工智能(Artificial Intelligence,简称 AI)术语,并将其定义为制造智能机器的科学与工程。至此,AI 诞生。

AI 诞生后,沿“弱人工智能—强人工智能—超人工智能”轨迹发展演化,充满挑战与创新。所谓弱人工智能(Weak AI),其实应称为狭义人工智能(Artificial Narrow Intelligence,简称 ANI),是专注于特定任务或领域(如自动翻译、自动驾驶、智慧健康、金融科技、智能化科研等)的人工智能,具有与人类相似的认知能力。所谓强人工智能(Strong AI),其实应称为通用人工智能(Artificial General Intelligence,简称 AGI),是适于任何领域或任务(包括科学创新、通用知识和社会交流等)的人工智能,能够像人类一样进行思考、计划、解决问题、抽象思维、理解复杂理念、快速学习和从经验中学习等操作,具有与人类同等的认知能力。所谓超人工智能(Artificial Super Intelligence,简称 ASI),是适于所有领域和任务的人工智能,具有全面超越人类的认知能力,可能最终会具有意识。

弱人工智能的典型代表之一是 DeepMind 公司推出的 AlphaFold。2020 年,DeepMind 公司的 Demis Hassabis 和 John M Jumper 提出了 AI 模型 AlphaFold2。该模型解决了一个已有 50 年历史的难题,能够预测大约两亿种已知蛋白质的复杂结构,并且已被全球 200 多万人使用。由于研发该模型的特殊贡献,Demis Hassabis 和 John M Jumper 二人获 2024 年诺贝尔化学奖。

强人工智能的典型代表之一是 OpenAI 公司发布 ChatGPT。2022 年,OpenAI 正式发布了 ChatGPT。针对所有领域或任何任务,ChatGPT 不仅能回答问题、创作文章、编程,甚至还可模仿人类的对话风格。

ChatGPT 几乎无所不能的回答能力,使人们对大语言模型的通用能力有了全新的认识。

超人工智能目前还处于理论推测阶段,其产品或模型短时还难以出现。

伴随以 AlphaFold 为典型代表的弱人工智能的广泛而高效地应用,以及 ChatGPT 为典型代表的生成式 AI 大模型的生成能力、迁移或推广能力、交互能力的不断强大,目前,我们已进入了弱人工智能(或狭义人工智能)的成熟应用阶段,已站在了通往强人工智能(或通用人工智能)的大门,已能感觉到超人工智能不久将会到来。

## 二、人工智能驱动人文社会科学研究新范式的形成

以 AlphaFold 和 ChatGPT 为代表的 AI 工具,展现出了卓越的解决复杂问题的能力,不仅驱动科学研究,也驱动人文社会科学研究新范式的形成。由 AI 驱动科学(AI for Science,简称 AI4S)的研究新范式,被人们称为科学研究第五范式<sup>[1]</sup>。同理,由 AI 驱动人文社会科学(AI for Social Science,简称 AI4SS)的研究新范式,被人们称为人文社会科学研究第五范式<sup>[2]</sup>。与科学研究第一范式(基于实验科学)、第二范式(基于理论科学)、第三范式(基于仿真科学)、第四范式(基于数据驱动的科学发现)<sup>[3]</sup>不同,人文社会科学研究第五范式(AI4SS)具有以下特征:(1)AI 全学科全过程赋能,支撑人文社会学科研究新视角;(2)人机智能融合,机器涌现的智能成为人文社会科学研究的组成部分;(3)概率统计模型在社会科学研究中发挥更大作用,组合爆炸(维度灾难)问题得以有效应对;(4)复杂理论受重视,AI 为人文社会科学研究构筑起多节点的复杂网络;(5)AI 虚拟现实促进人文社会科学理论与实践结合;(6)以大模型为特征的科研大平台成为人文社会科学研究的必需<sup>[4]</sup>。随着数据可获得性不断提高、算法技术不断更新以及计算能力的不断增强,AI4SS 也正在像 AI4S 驱动科学一样,全方位全过程地驱动人文社会科学的创新发展<sup>[4][5]</sup>。

### (一)AI 全学科全过程赋能,支撑人文社会学科研究新视角

以 ChatGPT、Llama、通义 Qwen 为代表的生成式人工智能(Generative artificial intelligence,简称 GAI),以其强大的内容生成能力、自动化处理特性和跨领域应用能力,全面赋能语言、文学、艺术、历史、哲学、经济、政治、法学、社会学、教育、心理、管理等人文社会科学,驱动人文社会科学间及人文科学与科学及技术间的交叉,全面应用于包括跨学科学习资源搭建、研究源流回溯、研究主题生成、研究趋势预测在内的科研全过程,将人文社会科学研究从线性视角扩大到非线性视角,为理解和解释人类社会现象提供了新的视野。

### (二)人机智能融合,机器涌现的智能成为人文社会科学研究的组成部分

在 AlphaFold 2 实现蛋白质结构预测中,机器猜想发挥了关键作用,这说明大规模的机器学习神经网络已涌现出某种程度的认知智能。所谓“机器涌现”,是指大语言模型或复杂系统,在达到一定规模或复杂性后,表现出一些意想不到的能力或特性,即“系统定量上的变化可以导致系统行为上的定性变化”。

目前,机器在人文社会科学领域涌现的智能形式多种多样,涵盖了文本与数据分析、预测与决策支持、艺术创作与文化创新以及人机交互与智能服务等多个方面。这些智能形式的应用,不仅提高了人文社会科学研究效率和质量,也为该领域的创新和发展提供了新的动力。

### (三)AI 概率统计模型在社会科学研究中发挥更大作用,组合爆炸问题得以有效应对

AI 在科研中的应用的主要目标之一是有效应对组合爆炸问题。所谓“组合爆炸”(数学上又称“维度灾难”),是指在解决某些问题时,可能的解决方案的数量随着问题规模的增大而呈指数级增长。社会科学中,比较典型的组合爆炸问题包括社会结构与关系网络、决策与路径选择、信息传播与谣言扩散、博弈与策略选择以及要素组合优化等问题。这些问题具有高度的复杂性和不确定性,给科研带来极大挑战。

解决社会科学研究中的组合爆炸问题的有效方法是追求概率统计意义上最终结果的可接受性。概率统计意义的正确性及确定性与计算程序的严格正确性是解决复杂问题的不同思路。人工智能研究的新近发展体现一种趋势:放弃绝对性,拥抱不确定性,即只求近似解或满足一定精度的解。这或许是这次人工智能“意外”取得成功的深层原因<sup>[1]</sup>。

为了有效应对社会科学研究中的组合爆炸问题的挑战,利用AI概率统计模型是必要的。所谓AI概率统计模型,是利用概率统计原理和方法,结合人工智能技术,对随机事件进行建模、预测和分析的模型。通过该模型,研究者可以更高效地处理和分析大规模数据集,更深入地理解和预测社会现象,为解决具有高度复杂性和不确定性的问题提供有力支持。

#### (四)复杂理论受重视, AI 为人文社会科学研究构筑起多节点的复杂网络

复杂网络由大量节点(代表个体或单位)和边(代表个体或单位之间的关系)构成。复杂网络理论是研究复杂网络的结构、性质、演化规律及其动力学行为的理论体系,在社交网络分析、组织行为研究、城市与人类行为研究、经济与市场行为研究等方面有广泛的应用场景。AI技术,特别是机器学习算法和深度学习算法,能够从海量的社会科学数据中提取有价值的信息,并根据这些信息构建出多节点的复杂网络。过去我们构建的网络是平面的,而现在人工智能将生成多维的、多节点的复杂网络。在这种复杂网络下,我们更能发现问题。

#### (五)AI 虚拟现实促进人文社会科学理论与实践结合

AI虚拟现实技术的大量应用,促进了人文社会科学理论与实践结合,尤其是应急响应、社会管理等社会科学。例如,就应急响应而言,AI虚拟现实技术通过模拟真实的环境和场景,使研究人员和应急响应人员能够在虚拟环境中进行实践操作。这种实践操作不仅能提高他们的技能水平,还能帮助他们更好地理解理论知识在实际应用中的具体表现。就社会管理而言,通过AI虚拟现实技术,政策制定者可以更加直观地了解政策实施后的实际效果和影响。这种直观的模拟和预测有助于政策制定者更加科学地制定和实施政策,提高政策的有效性和针对性。

#### (六)以大模型为特征的科研大平台成为人文社会科学研究必需

数字及智能时代,智能设备为人文社会科学研究提供了丰富多样的数据(如社交媒体数据、生理与行为数据、心理与行为数据、环境数据、地理位置数据、历史文献数据等),文献信息数据库为人文社会科学研究提供了海量的数据、信息和文献,工具研发者为人文社会科学研究提供了大量的工具(如文献管理工具、数据分析工具、学术资源检索工具、社交网络分析工具、思维导图工具等),人文社会科学研究数据化及工具化成为趋势。

AI大模型,因其强大知识存储及生成知识能力,已成为人文社会科学研究数据化及工具化的理想平台。它可以赋能数据、信息、文献的收集、整理和分析,将分散的碎片化知识整合起来,形成知识体系。可以赋能数据分析,发现不同数据之间的关联性和规律性,从而揭示出隐藏在数据背后的逻辑关系。它势必成为人文社会科学研究必需。

### 三、人工智能赋能人文社会科学研究中的选题及学术史梳理

人工智能驱动的人文社会科学研究(AI4SS)引发人文社会科学领域的革命,在人文社会科学研究中赋能选题与学术史梳理,全面驱动假设生成及假设论证方法变革。本部分首先讨论AI赋能人文社会科学研究中的选题及学术史梳理。

选题,即选择研究题目,是人文社会科学研究的第一个环节。无论是项目申报,还是论文著作撰写,都要确定好选题,“好的选题等于成功了一半”。一般而言,确定选题需要依次完成五件事——确定研究方

向、明晰/细化研究问题(包括明确研究对象、侧重点、时间范围)、确定研究主题、确定研究内容、确定研究题目,而这五件事的完成,都有赖于通过获取文献、阅读文献、判别文献质量等来进行学术史梳理。人工智能,特别是以 ChatGPT 为代表的生成式大语言模型,由于其强大的自然语言处理能力,广泛的知识覆盖范围,非凡的辅助理解、提取关键信息和批量处理文献的能力,出色的智能推荐功能,为文献检索、文献阅读和文献质量判断带来了新的方式方法,为学术史梳理提供了便捷,为选题提供了支撑。

### (一) 文献检索

生成式人工智能为学术资源检索提供了面向问题检索的能力,跳出了关键词、主题词等检索框架。所谓面向问题检索,就是通过与 AI 进行对话,研究者将自己的检索问题与需求告知 AI, AI 通过自然语言处理技术理解研究者的真实需求,自动判断并选取匹配出最重要最相关的文献(甚至是知识或答案)提供给用户的这样一个人机交互的过程。例如,我们检索《实践是检验真理的唯一标准》一文的作者是谁?我们用生成式人工智能(文心一言),得到的答案是胡福明(南京大学)、孙长江(首都师范大学),这就是面向问题检索。而我们用文献数据库如《光明日报》(1949—2018)全文检索数据库检索,得到的结果只能是简单的结果:“特约评论员”(特约评论员.实践是检验真理的唯一标准 [N].光明日报,1978.5.11)。面向问题检索具有两大特征:一是自然语言、面向目标、针对问题,借助大模型、大语言实现语义推理的智能检索;二是针对检索结果, AI 可判断结果与问题的密切程度。

实现面向问题检索的先决条件是要对文献数据相互间的语义关系进行标注。通过标注,才能使我们在不知道被检索对象的准确检索词(如作者、题名、关键词等)的情况下,通过分析被检索对象相关数据相互间的语义关系,检索并获取到最终需要的结果。

以前,在没有计算机的时候,我们需要到图书馆利用题名目录、作者目录、分类目录、主题目录等工具检索文献,但以这种方式得到的结果不是我们最需要的目标,而仅仅是文献资源。而现在及未来的检索,基于 AI 大模型的检索,由于具有语义检索环境,不再受词汇及语种的检索限制,让我们可用自然语言(甚至可以不用打字)直接检索到目标。

### (二) 文献阅读

AI 可赋能学术阅读,提高阅读效率,帮助读者省力阅读、聚焦阅读。所谓省力阅读、聚焦阅读,就是通过人工智能把文献的关键内容直接给阅读者,让阅读者省力并聚焦其关注的问题。一篇论文,有 80% 可以不用看。但阅读者怎么才能够快捷看到那 20% 的内容呢?是不是必须把 80% 看完后才能看到那 20% 呢?所以,阅读者可以用 AI 省力阅读、聚焦阅读,实现阅读的人机交互,对想要阅读的内容进行摘编。

通过 AI 辅助阅读,阅读者既可以分析出文章的重要观点、重要理论、方法和要解决的问题,也可以根据文章的内容写出自己的提示假设和对原文假设进行的思考、评价、完善或更新。就特定选题而言,阅读者越多地汇集文章提示假设,就越能提出自己的新假设。

为方便文献阅读理解, AI 应提供一个辅助生成阅读卡片的系统。利用该系统,阅读者可将自己的阅读体会记录在卡片上。上世纪七八十年代,很多文科学者,每读一篇文章就做一张卡片。本文作者在那时也是每读一篇文章做一张卡片,所做的卡片累积很多,一张一张,薄薄的。还做导卡,指引这些卡片是什么内容。还按主题归类,将阅读的与该主题相关的文章的笔记汇集起来。阅读笔记的内容包括文章讲的是什么,阅读者的感想是什么,文章给了阅读者什么启示。后来,在做文献阅读软件的时候,我们就据此提出思路,对该软件的功能提出要求,比如,在电脑显示屏上,要求做到一边是一篇文章原文,另一边

就是用 AI 析出的这篇文章的主要观点、理论、方法、解决了什么问题, AI 提出的假设, 以及读者在阅读 AI 析出内容后自己的感言及假设。有了以上这些阅读笔记的积累, 读者梳理选题学术史以及把握选题研究进展的效率就高多了。

怎么利用读者的读书笔记生成学术史梳理及研究进展的内容呢? 一般而言, 在当前 AI 支持下, 读者通过阅读卡片积累的信息很多, 有摘要信息, 有人工智能的观点, 有人工智能的预判假设信息, 还有读者自己的体会及相关标签。当某个领域(或主题)的文献阅读笔记达到一定数量(比如 50 篇)时, 读者自己就可自动地通过 AI 生成一个选题引言、学术史梳理及研究进展了。

通过 AI 辅助生成选题以及做好学术史梳理和研究进展把握, 建立读者个人阅读心得库是关键。读者无论从文献数据库(比如知网)下载多少篇文献, 原则上都可以建立自己的阅读心得库。现有文献阅读心得库软件会自动给读者创建一个目录, 方便直接调用。

### (三) 论文质量判断

利用传统的关键词(词)在文献数据库中检索文献, 或在网络上搜索文献, 得到的检索结果往往数量大, 质量无保障。如果研究者逐篇阅读, 必然耗费大量时间与精力。因此, 在选题及学术史梳理过程中, 有必要利用 AI 助力论文质量判断。

判断论文质量的主要指标是创新程度、前沿性及学术规范性。论文成果的创新程度及前沿性判断, 以前是靠小同行评审, 但该方法费时的同时花费也大。如今 AI 已能帮助分析出文章解决了什么重大理论与现实问题, 提出了什么新理论等内容, 这为评价论文创新程度及前沿性提供了可能。论文学术规范性评价, 其主要内容包括论文研究目的与问题的明确性、论文研究方法的科学性与可行性、论文文献综述的完整性与相关性、论文研究结果与讨论的清晰性与深入性、论文结论的准确性与前瞻性、论文格式的规范性以及论文作者的学术道德及科研伦理。如今的 AI, 由于其具有的强大的知识库功能, 已能对这些内容进行深入分析了, 这为评价论文学术规范提供了支撑。

## 四、人工智能赋能人文社会科学研究中的假设生成

假设是研究者对某种行为、现象或事件做出的一种合理的、尝试性并有待检验的解释。假设生成是研究者创造性思维的过程, 有假设, 才可能有新发现。

以前, 研究人员是从现有的理论、数据中汲取灵感, 采用头脑风暴法形成设想。然而, 随着 AI 的不断发展以及文献信息量的急剧增长, 研究人员需要以更快、更有效的方法来生成可靠的假设。目前, 来自人文社会科学不同领域的研究人员试图使用大语言模型来帮助其完成这项任务, 例如, Yang Jeong Park、Daniel Kaplan 及 Zhichu Ren 等学者<sup>[6]</sup>研究了大语言模型能否像人类那样创造性地进行假设生成这一问题。他们使用 GPT-4 生成科学假设, 并得出研究结论: 虽然错误率很高, 但生成式人工智能似乎能够有效构建大量的科学知识并提供有趣且可测试的假设。他们认为, 在未来的科学研究中, 科学家们可能会与大量的“假设机器”进行协同工作。他们还认为, 这种协同工作将会面临两大挑战: 一是自动化实验的挑战, 即如何确保实验过程的准确性和可靠性, 以及实验结果的有效性; 二是对抗性同行评审(Adversarial Peer Review)的挑战, 即在同行评审过程中, 评审者可能会以更为挑剔或对抗性的态度来审视和评估这些由“假设机器”辅助产生的科学成果。

Zonglin Yang、Xinya Du 及 Junxian Li 等学者<sup>[7]</sup>对大语言模型赋能自动开放领域的科学假设进行了研究。他们指出, 以前, 科学假设被认为是科学家在观察世界并试图提出假设来解释这些观察时的主要推理类型, 其研究受两方面限制: (1) 数据集中的观察注解是精心手动挑选的句子(导致领域封闭); (2) 真实

假设大多为常识知识,使得任务相对不那么具有挑战性。为突破这两方面的限制,他们提供了一个支撑社会科学学术假设生成的数据集,并试图最终创建一个能够仅凭原始网络语料库自动生成有效、新颖且有帮助的科学假设系统。与以前的设置不同,新数据集需要使用开放领域数据(原始网络语料库)作为观察,并提出人类前所未有的假设。他们的研究成果是,开发了一个科学假设系统多模块框架。该框架用三种不同的反馈机制提高性能,在基于 GPT-4 和基于专家的评估方面表现出了卓越的性能。他们的研究表明,LLMs 能够生成新颖的“文献中不存在的”和有效的“反映现实的”科学假设。

S Banker 等学者<sup>[8]</sup>采用大型自然语言模型(LLM)辅助社会心理学假设生成。他们试验了两种研究方法。第一种方法,他们用过去 55 年间在 50 多家社会心理学期刊上发表的数千篇论文摘要以及预印本存储库(PsyArXiv)中的资料,对 GPT-3 语言模型进行了微调。参与该研究的社会心理学专家认为,在清晰度、原创性和影响力三个方面,模型生成的假设的质量和人类生成的类似。第二种方法,他们未对 GPT-4 进行微调,而是直接生成假设。参与该研究的社会心理学专家认为,在清晰度、原创性、影响力、合理性和相关性五个方面,这些直接生成的假设的质量比人为产生的更高。这两种方法得出的研究结果表明,LLM 生成的假设与人类研究者撰写的假设不相上下,甚至更胜一筹。

清华大学童松(Song Tong)等学者<sup>[9]</sup>,利用因果知识图谱与大型语言模型(LLM)之间的协同,提出心理学领域假设生成新方法。他们利用大型语言模型分析了 43 312 篇心理学文章,提取了因果关系对。他们通过应用链接预测算法,生成了 130 个以“幸福感”为重点的潜在心理学假设,并将其与博士生构思的研究思路以及仅由大型语言模型产生的研究思路进行了比较。他们发现,结合大语言模型和因果图的方法提出的假设在新颖性方面与专家不谋而合,明显优于仅由大型语言模型生成的假设。该项研究的结果表明,将大型语言模型与因果知识图谱等机器学习技术相结合,可以彻底改变心理学领域的自动化发现过程,从海量文献中提取出新颖见解。该项研究处于心理学与人工智能的交汇点,为心理学研究中的数据驱动假设生成提供了新的方法。

中国科学院大学徐若曦(Ruoxi Xu)等学者<sup>[10]</sup>在完成 AI 在假设生成方面的应用的综述后指出,目前,大型语言模型在研究初期只能作为辅助和启发工具,并且存在以下局限:第一,大语言模型缺乏对输出内容有效性的理解,导致信息捏造或不正确,从而误导用户;第二,对提问提示高度敏感,导致研究人员在提问提示设计上投入大量精力,但产出结果却不确定;第三,上下文长度有限,难以处理长篇幅和多篇文章。他们认为,大语言模型应用于假设生成的未来方向包括:(1)通过检索增强技术或领域特定训练来整合专业知识;(2)开发高回报的提示策略;(3)扩大大型语言模型的上下文窗口。通过允许模型考虑更广泛的上下文,使其能够访问更全面的信息,从而产生更稳健且富有洞察力的假设。

## 五、人工智能赋能人文社会科学研究中的假设论证

就一个特定的研究课题而言,假设(研究主题)和分假设(分论题)一旦确立,紧接着的工作就是开展假设论证。假设论证是收集和分析数据以验证假设为真或为伪的过程。当对一系列假设完成了验证,新的理论可能就形成了。目前,人们将人文社会科学研究中的假设论证方法分为理论论证法和实证论证法两类,其中实证论证法又分为实验论证法和非实验论证法两种。非实验论证法又可再分为调研论证法(统计数据调研论证法、实地调研论证法)及无干扰研究论证法(文本分析论证法、现有统计数据论证法、事件研究论证法)。下面,特就人工智能在这些假设论证方法中的应用进行论述。

### (一) 理论论证法

理论论证法是通过逻辑推理和理论分析来验证假设的方法。它主要依赖于已有的理论框架、数学

模型或逻辑体系,通过演绎或归纳的方式对假设进行验证。它的特点是强调逻辑性和理论性,不依赖于具体的实验数据或实地观察,更多地依赖于学者的理论素养和逻辑推理能力。生成式 AI 大模型由于具有强大的知识库能力、内容生成能力、计算能力和逻辑推理能力,可以为假设的理论论证提供以下辅助。

### 1. 文本生成与理论阐述

基于已有的理论框架和文献,生成与假设相关的文本内容,辅助研究者更清晰地阐述假设,并构建理论论证的逻辑链条。例如,研究者可以利用大模型生成与特定理论相关的概念辨析、理论关系阐述等文本内容。

### 2. 数据挖掘与知识发现

处理和分析大规模的数据集,从中挖掘出潜在的知识 and 规律。通过数据挖掘,研究者可以发现与假设相关的新的证据或观点,从而增强理论论证的说服力。

### 3. 逻辑推理与假设验证

对假设进行演绎或归纳式的验证。虽然这一过程仍需要人类的参与和校验,但大模型可以提供初步的逻辑推理结果,帮助研究者发现潜在的问题或矛盾,并进一步完善理论论证。

## (二) 实证论证法

### 1. 实验论证法

实验论证法是通过设计实验来验证假设的方法,具有高度的可控性和可重复性,能够准确地揭示因果关系,适用于那些可以通过实验来操纵和观察的社会现象。目前, AI 对实验论证法的辅助主要体现在实验助手及实验模拟两个方面。

#### (1) 实验助手

实验助手是指在实验中使用 AI 大模型来辅助完成一些简单但劳动密集型的任务,这些任务以前通常由研究人员手动完成。例如,可以利用 AI 协助研究人员通过反馈来迭代创建假设场景,以增强实验条件的外部有效性和可比性<sup>[4]</sup>。此外,大型语言模型能够合成实验所需的信息,从而避免了使用现实生活中的信息。这样做,有利于保护其信息可能被用于研究的个人的隐私。

#### (2) 实验模拟

实验模拟主要通过模型、软件或虚拟环境来模仿真实的实验过程和结果。在模拟实验中,大型语言模型通常被用作人类行为的代理,它们不仅能够模拟人类行为,还能在复杂的系统中揭示出有趣的行为模式和交互动态。这种方法为社会科学研究提供了新的视角和工具,使得研究人员能够在不直接涉及真实人类参与者的情况下,探索和理解复杂的社会现象。

### 2. 非实验论证法

#### (1) 调研论证法

调研论证法是通过收集和分析大量数据来验证假设的方法,通常依赖于问卷调查、实地调查、人口普查等数据来源,是人文社会科学研究中的一种常用方法。调研论证法通过书面问卷或正式访谈来收集目标群体的信仰、观点、特征以及过去或现在的行为信息。当前, AI 语言大模型已应用于调研论证法中的抽样、观测及分析。

##### ① 抽样

抽样是从组成总体的所有元素的集合中,按一定的方式选择或抽取一部分元素的过程,涉及从人类群体中选取具有代表性的样本。 AI 大模型为抽样提供了一种新颖的选择,可以作为特定人类群体的代

理,或者说可将AI大模型的庞大训练数据库作为研究样本(“硅基样本”)<sup>[11]</sup>。

### ②问卷设计与调查测评

在调研论证法中,问卷设计与调查测评是获取数据的重要环节。目前,利用AI语言大模型不仅可以辅助设计问卷、量表或访谈提纲,还可促进问卷调查实现“封闭式评分量表—开放式回答问卷—更自然的交互式访谈”这一转化。从O.N.Kjell、S Sikström、K Kjell等学者<sup>[12]</sup>的研究成果看,在心理调查方面,使用基于自然语言的开放式问卷进行的调查结果与使用评分量表进行的心理调查的结果相比,前者实现的准确性要么超过后者(这通常被视为理论上限),要么与之相当。从O Kjell、K Kjell、H.A Schwartz等学者<sup>[13]</sup>的研究成果看,在心理健康评估方面,基于人工智能的大语言模型已能应用于调查测评,且以后的人工智能大语言模型在调查测评过程中会提出更精细粒度的自动交互式访谈。

### ③数据分析

数据分析是对调查测评到的数据进行整理、加工、解释和应用的过程,其目的在于从这些数据中发现规律、趋势和模式。大型语言模型可用于分析这些数据。例如,Yang、Ji、Zhang、Xie等学者<sup>[14]</sup>利用大型语言模型,特别是ChatGPT,进行心理健康分析,认为大型语言模型在提高心理健康分析可解释性方面具有巨大潜力。

## (2)无干扰研究论证法

无干扰研究论证法是在不干扰研究对象自然状态的情况下收集和分析数据来验证假设的方法,它主要依赖于已有的文本资料、现有统计数据或历史事件等,它具有客观性和真实性。将AI大语言模型应用于该论证法中,利于消除数据资源的限制,便于跨学科分析问题并提出解决问题的方案,可以通过文本输入及直接交互,降低以往论证分析难度。下面探讨AI大语言模型在文本分析论证法和现有统计数据论证法中所起的作用。

### ①文本分析论证法

文本分析论证法是通过分析已有的文本资料(如历史文献、政策文件、新闻报道等)来验证假设。它适用于那些可以通过文本分析来揭示的社会现象。目前,人们可以利用AI大语言模型执行文本检测及分类标注任务,如完成文本情感分析以识别和提取文本中的情感态度,完成文本立场检测以确定作者在文本中表达的政治、社会或文化立场,完成文本仇恨言论检测以识别和过滤文本中可能包含仇恨言论的单词、短语或句子,完成文本虚假信息检测以识别和过滤文本中可能包含虚假信息的单词、短语或句子;还可以利用AI大模型执行文本生成任务,如完成自然语言描述或解释以及未来预测。

### ②现有统计数据论证法

现有统计数据论证法是利用已有的统计数据(如人口普查数据、经济统计数据等)来验证假设,适用于通过统计数据来揭示社会现象的本质。目前,AI大语言模型可用于描述样本的特征或变量之间的关系,或对因果过程进行推断,这被称为描述性和推断性分析。还可以根据历史数据推断未来的趋势和变化,这被称为预测性分析。

## 结语:面向人文社会科学研究开源开放大模型建设

AI为人文社会科学研究带来了新的研究范式,为人文社会科学研究注入了新的学术史梳理、假设生成及假设论证方法。为促进AI赋能研究方法变革,当前和今后,人文社会科学界应加快面向人文社会科学研究开源开放大模型建设。

截至目前,Meta和谷歌推出了开源开放大模型Llama和Gemma<sup>[4]</sup>,清华大学和北京智源人工智能研

究院推出了开源开放大模型 ChatGLM-6B(<https://github.com/THUDM/ChatGLM-6B>)和悟道·天鹰(Aquila)(<https://model.baai.ac.cn/models>)。这些大模型向公众公开模型源代码、数据以及权重等信息,促进了其使用的平民化。但对我国社会科学研究而言,目前已有的这些开源开放大模型,普遍缺乏针对性,所完成的预训练数据尚未全面覆盖各主要语种文献及各类型文献(特别是图书、研究报告、古籍、数据集)。

要完成对多语种文献数据和多类型文献数据的预训练,建成针对性强的人文社会科学研究开源开放大模型,工作量特别大,专业性特别强,因此,人文社会科学界应共同努力,大力推进面向人文社会科学研究开源开放大模型建设,为我国人文社会科学研究提供良好数字智能环境。

### 参考文献:

- [1] 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.
- [2] 庞珣. 人工智能赋能社会科学研究探析——生成式行动者、复杂因果分析与人机科研协同[J]. 世界经济与政治, 2024(7): 3-30.
- [3] TONY Hey, STEWART Tansley, KRISTIN Tolle. 第四范式: 数据密集型科学发现[M]. 潘教峰, 等译. 北京: 科学出版社, 2012.
- [4] BAIL C A. Can generative AI improve social science?[J]. Proceedings of the national academy of sciences of the United States of America, 2024, 121(21): e2314021121.
- [5] 苏新宁, 何一民, 杨红艳, 等. AI赋能哲学社会科学研究、传播及评价[J]. 西华大学学报(哲学社会科学版), 2024(6): 1-12.
- [6] PARK Y J, KAPLAN D, REN Z C, et al. Can ChatGPT be used to generate scientific hypotheses?[EB/OL]. [2024-12-06]. <http://arxiv.org/abs/2304.12208>.
- [7] YANG Z L, DU X Y, LI J X, et al. Large language models for automated open-domain scientific hypotheses discovery[EB/OL]. [2024-12-05]. <http://arxiv.org/pdf/2309.02726v2>.
- [8] BANKER S, CHATTERJEE P, MISHRA H, et al. Machine-assisted social psychology hypothesis generation[J]. *The American psychologist*, 2024, 79(6): 789-797.
- [9] TONG S, MAO K, HUANG Z, et al. Automating psychological hypothesis generation with AI: When large language models meet causal graph[J]. *Humanities and social sciences communications*, 2024, 11(1): 896.
- [10] XU R X, SUN Y F, REN M J, et al. AI for social science and social science of AI: A survey[EB/OL]. [2024-12-06]. <http://arxiv.org/abs/2401.11839>.
- [11] 殷杰. 生成式人工智能的主体性问题[J]. 中国社会科学, 2024(8): 124-145.
- [12] KJELL O N, SIKSTRÖM K, SCHWARTZ H A. Natural language analyzed with AI-based transformers predict traditional subjective well-being measures approaching the theoretical upper limits in accuracy[J]. *Scientific reports*, 2022, 12(1): 1-9.
- [13] KJELL O, KJELL K, SCHWARTZ H A. AI-based large language models are ready to transform psychological health assessment[J/OL]. [2024-12-06]. <https://doi.org/10.31234/osf.io/yfd8g>.
- [14] YANG K L, JI S X, ZHANG T L, et al. On the evaluations of ChatGPT and emotion-enhanced prompting for mental health analysis[EB/OL]. [2024-12-06]. <https://arxiv.org/abs/2304.03347v1>.

[责任编辑 李秀燕]