

# 人工智能在科学发现中的角色定位与认识论边界

韩建伟

青海高等职业技术学院, 青海海东 810799

**【摘要】**科学发现的逻辑与可行性一直是哲学与科学界争议的核心议题。随着人工智能技术的迭代,机器学习、神经网络等方法逐步渗透到科学研究的核心环节,为科学发现提供了全新路径。本文基于科学哲学、认知科学与人工智能交叉视角,结合相关研究成果,系统剖析人工智能在科学发现中的三重角色:数据处理与模式识别的高效工具、科学概念与理论的辅助建构者、科学发现过程的模拟与复刻者。同时,从数据依赖、因果推理、理论创新及心理社会因素四个维度,厘清人工智能介入科学发现的认识论边界。研究表明,人工智能能够拓展科学发现的广度与效率,但无法替代人类在理论创新、价值判断及复杂因果建模中的核心作用,跨学科融合是突破其边界的关键路径。

**【关键词】**人工智能;科学发现;认识论边界

## 一、引言

科学发现是否存在可遵循的逻辑,能否通过系统性方法实现自动发现,这一问题困扰了哲学家与科学家半个多世纪。逻辑实证主义将发现与辩护割裂,认为科学发现缺乏普遍逻辑;而人工智能先驱赫伯特·西蒙则主张,科学发现是特殊的问题求解过程,可通过计算机程序模拟。早期人工智能依赖规则推理与数据挖掘,虽在特定领域实现了经验规律的再发现,却因局限于“曲线拟合”或依赖先验知识而备受争议。

随着大数据与深度学习的崛起,机器学习不再局限于局部问题的降维处理,开始具备从原始数据中提炼科学概念、构建简单理论的能力。从物理学概念的自动识别到化学反应路径的发现,人工智能正逐步改变科学研究的传统模式。但与此同时,其在真实科学场景中的适用性、理论创新的深度及对科学本质的把握等问题,也引发了新的认识论思考。

本文结合科学哲学理论与人工智能研究案例,聚焦两个核心问题:人工智能在科学发现中究竟扮演何种角色?其能力边界与认识论局限何在?通过整合相关文献的核心观点与实证研究,试图为理解人工智能与科学发现的关系提供系统性框架。

## 二、人工智能在科学发现中的角色定位

人工智能对科学发现的介入,并非简单替代人类研究者,而是通过技术优势弥补传统研究方法的不足,形成“工具赋能—过程模拟—理论辅助”的多层次角色体系。

### (一) 数据处理与模式识别的高效工具

现代科学已进入大数据时代,天文学、分子生物学等领域的观测设备每日产生海量数据,稀有科

学事件的捕捉与分析远超人类经验处理能力。人工智能凭借分布式处理框架与高效算法，成为科学数据处理的核心工具。

在数据实时分析层面，针对科学数据流的“快、准、全”需求，人工智能构建了在线-离线交互反馈机制：离线层通过海量数据训练高精度模型，涵盖数据全局特征；实时层借助模型迁移实现快速异常检测，同时通过反馈迭代持续优化算法。例如，时域天文学中，人工智能系统可处理 TB 量级的每日观测数据，从高噪声中识别短时标稀有事件，其报警率控制在十万分之一以下，远超人工处理效率。

在模式挖掘层面，机器学习能够从高维数据中提取隐藏关联。苏黎世联邦理工学院的“科学网”（SciNet）通过表征学习，将高维观测数据压缩为低维表征，成功从阻尼摆运动数据中识别出弹性系数与阻尼系数概念，从天文观测数据中重构出日心说系统。这种模式识别能力不仅突破了人类认知的局限，还能规避传统研究中可能存在的主观偏见。

## （二）科学概念与理论的辅助建构者

科学发现的核心在于从现象中提炼概念、从局部规律中整合为普适理论。早期人工智能的理论建构局限于经验规律的再发现，而现代机器学习已具备初步的概念生成与理论整合能力。

在概念发现层面，人工智能通过无监督学习实现从数据到概念的跃迁。清华大学与加州大学圣迭戈分校的研究团队构建 RNN 神经网络，从基本粒子的势能与概率密度数据中，自主学习到薛定谔方程，验证了人类构建的物理理论是数据所能支撑的最优模型。BACON 程序通过选择性搜索启发式，不仅重新发现了波义耳定律、开普勒第三定律等经典规律，还自主提炼出比热、电压等核心科学概念。这些案例表明，在缺乏先验知识的情况下，人工智能能够通过数据驱动的方式生成科学概念。

在理论建构层面，麻省理工学院提出的“AI 物理学家”框架，整合了分治算法、奥卡姆剃刀、统一理论与终生学习四大模块。该框架先通过分治算法挖掘不同领域的局部理论，再以最小描述复杂性原则筛选简洁理论，最终实现局部理论的整合与通用化。在带电双摆与多环境穿梭物体的测试中，模型成功识别出引力、简谐运动等核心规律，展示了理论整合的潜力。此外，KEKADA 程序通过模拟长期实验室工作，结合意外现象反应机制与可信度评估，不仅复刻了尿素合成的化学反应路径发现过程，还能提出新的实验假设，体现了人工智能在理论建构中的辅助价值。

## （三）科学发现过程的模拟与复刻者

科学发现不仅是结果的产出，更是过程的展开。人工智能通过模拟人类科学家的认知过程与研究方法，为理解科学发现的本质提供了全新视角。

西蒙将科学发现界定为问题解决的特殊形式，主张通过启发式搜索与信息处理模型模拟这一过程。其团队开发的“逻辑理论家”程序，成功复刻了罗素《数学原理》中的部分定理证明，成为人工智能模拟科学发现的里程碑。在此基础上，MECHEM 程序通过启发式策略，不仅重新发现了克雷布斯提出的尿素合成路径，还拓展出全新的有机物生成路径，证明科学发现的核心机制可通过算法复刻。

人工智能对认知过程的模拟还延伸到直觉与灵感等非逻辑因素。西蒙认为，直觉是“不经有意识推理而快速获得解答的能力”，灵感则是启发式搜索的副产品，可通过“生成-检验”机制模拟。BACON 程序的选择性搜索策略——从简单函数入手、聚焦变量单调关系、控制单一变量变化——正是对人类直觉思维的算法化复刻。而 KEKADA 程序引入的意外现象反应能力，能够模拟科学家对偶然发现的敏

感度，进一步贴近真实科学发现的不确定性。

### 三、人工智能在科学发现中的认识论边界

人工智能在拓展科学发现效率的同时，也受限于技术特性与科学本质的双重约束，其认识论边界主要体现在以下四个维度。

#### （一）数据依赖的局限性：模拟数据与真实场景的割裂

当前人工智能的科学发现研究多基于模拟数据展开。无论是“科学网”对物理概念的识别，还是“AI 物理学家”的理论构建，其训练数据均是通过已知理论反推生成的模拟数据，即使加入高斯噪声，本质上仍属于对已知规律的反向验证。而真实科学发现的核心挑战在于，研究者事先并不知道哪些数据相关，需要在有限资源下通过可控实验试错探索。

苏黎世联邦理工学院的日心说模拟，使用的是基于现代观测数据倒推的哥白尼时代数据，而非当时真实的观测记录与技术条件；KEKADA 程序虽模拟了尿素合成的发现过程，但依赖于对历史实验日记的结构化处理，未能还原真实研究中数据采集的复杂性与偶然性。此外，科学数据的长期存储与高效利用仍是未解决的难题，大量历史数据因管理不善而价值递减，无法为人工智能提供完整的训练基础。这种模拟数据与真实场景的割裂，使得人工智能难以应对真实科学发现中的数据稀缺、噪声干扰与变量冗余等问题。

#### （二）因果推理的短板：相关性与因果性的认知鸿沟

科学发现的核心目标是揭示事物间的因果关系，而非仅仅识别相关性。计算机科学家朱迪·珀尔指出，所有机器学习本质上都是“曲线拟合”，专注于挖掘数据中的相关性模式。尽管部分物理理论可通过统计拟合获得，但大部分自然科学与社会科学的研究对象都涉及复杂因果关系，需要超越数据表面的关联。

当前人工智能的因果建模能力仍显不足。例如，机器学习能够发现“地面湿”与“下雨”的相关性，却无法仅凭数据区分“下雨导致地湿”与“洒水导致地湿”的因果差异。“AI 物理学家”虽能识别物理规律，却难以处理多因素叠加的复杂因果场景——真实科学研究中，如带电物体在磁场中的运动，重力场与电磁场的影响同时存在，而模型仅能区分时间上顺序排列的单一环境。蔡曙山的研究也表明，因果推理涉及人类的经验积累与心理加工，即使是演绎推理，也会受到工作记忆、认知负荷等心理因素影响，这些都是当前人工智能难以模拟的。

#### （三）理论创新的边界：常规科学与科学革命的分野

库恩将科学发展划分为常规科学与科学革命两个阶段，常规科学是范式内的解题活动，而科学革命则涉及范式转换与概念重构。人工智能目前的理论创新能力，仍局限于常规科学层面。

现有人工智能系统的理论建构，本质上是对已有数据的归纳整合与局部优化。BACON 程序的规律发现、MECHEM 的反应路径推导，均是在既定理论框架内的再发现；“AI 物理学家”的理论统一，也依赖于局部理论的相似性匹配。这些工作都未能触及科学革命的核心——概念体系的根本性变革。例如，相对论对经典力学中“时间”“空间”概念的重构，量子力学对传统因果观的颠覆，均涉及人类

对世界本质的全新认知，需要超越数据的抽象思维与价值判断。而人工智能缺乏对科学概念内涵变化的敏感性，无法模拟这种范式转换过程。正如塞尔通过“中文屋论证”所揭示的，人工智能的符号处理缺乏意向性与语义理解，难以真正把握科学概念的本质。

#### （四）心理社会因素的缺失：科学发现的人文维度

科学发现不仅是逻辑与数据的产物，还深受研究者的经验、直觉、情感及社会文化背景影响。蔡曙山提出“所有逻辑都是心理逻辑”，认为溯因、归纳等推理过程均交织着经验与心理因素。沃森实验表明，人类推理会受到日常经验、认知偏好的影响，即使是演绎推理，也存在明显的个体差异。这些心理社会因素构成了科学发现的人文维度，而这正是人工智能的先天短板。

人工智能无法复制人类科学家的学术积淀与直觉洞察。查德威克之所以能从实验数据中迅速识别中子，源于其对卢瑟福中子假说的预先了解；西蒙构建“逻辑理论家”的灵感，来自对人类几何解题思维的长期观察。这些基于经验与直觉的判断，依赖于人类大脑的综合认知能力，而非单纯的算法运算。此外，科学发现的社会性与长期性也超出了人工智能的模拟范围，真实科学研究中的合作交流、学术争议、资源分配等社会因素，共同塑造着发现过程，而当前人工智能系统多为封闭的算法模型，缺乏与社会环境的交互。

## 四、结论与展望

人工智能在科学发现中扮演着三重递进的角色：作为数据处理工具，它突破了人类的认知局限，实现了海量数据中稀有模式的高效识别；作为理论建构辅助者，它能够从原始数据中提炼科学概念，整合局部规律为通用理论；作为过程模拟者，它复刻了科学发现的核心机制，为理解认知过程提供了实证支持。这些角色定位表明，人工智能已成为科学发现的重要赋能工具，拓展了科学研究的广度与效率。

但与此同时，人工智能的认识论边界也清晰可见：数据依赖使其难以脱离模拟场景走向真实发现，因果推理短板限制了其对科学本质的把握，理论创新局限于常规科学层面，心理社会因素的缺失使其无法复制科学发现的人文维度。这些边界并非技术迭代所能单独突破，而需要科学哲学、认知科学与人工智能的跨学科融合。

未来，突破人工智能认识论边界的路径可能包括三个方向：一是构建真实场景的数据集，结合科学史研究还原历史数据与研究条件，增强模型的现实适应性；二是发展因果推理算法，将科学哲学的因果理论与机器学习结合，提升模型对因果关系的建模能力；三是融入心理社会因素，借鉴认知科学的研究成果，构建更贴近人类认知过程的智能系统。正如孟小峰所提出的，科学数据智能需要建立“智能分析—知识融合—数据存储”的一体化框架，而人工智能与科学发现的深度融合，也需要建立“技术工具—哲学反思—认知模拟”的跨学科体系。

科学发现的终极目标是揭示世界的本质规律，这一过程既需要人工智能的技术赋能，也离不开人类的理论思辨与价值判断。人工智能是拓展科学发现边界的强大工具，但并非替代人类的“自主科学家”。在技术与人文的互补中，科学发现将走向更高效、更深刻的新境界。

#### 参考文献

- [1] 王东. 机器学习与科学发现的逻辑刍议[J]. 北京航空航天大学学报(社会科学)

- 版), 2021, 34(2): 99-104.
- [2] 蔡曙山. 科学发现的心理逻辑模型 [J]. 科学通报, 2013, 58(34): 3530-3543.
- [3] 李西君, 齐爱花. 科学发现中直觉与灵感的模拟性 [J]. 宁波广播电视大学学报, 2007, 5(1): 105-108.
- [4] 王天珍. 计算机模拟科学发现——西蒙 (Simon) 教授近年来的一些研究工作简介 [J]. 江汉大学学报 (社会科学版), 1995(4): 95-99.
- [5] 孟小峰. 科学数据智能: 人工智能在科学发现中的机遇与挑战 [J]. 中国科学基金, 2021, 35(3): 419-425.

# The Role and Epistemological Boundaries of Artificial Intelligence in Scientific Discovery

HAN Jianwei

Qinghai Higher Vocational & Technical Institute, Haidong, Qinghai 810799, China

---

**Abstract:** The logic and feasibility of scientific discovery have long been core topics of debate in philosophy and science. With the iterative advancement of artificial intelligence technology, methods such as machine learning and neural networks have gradually permeated the core processes of scientific research, offering new pathways for scientific discovery. From an interdisciplinary perspective integrating philosophy of science, cognitive science, and artificial intelligence, and drawing on relevant research, this paper systematically analyzes the threefold role of AI in scientific discovery: an efficient tool for data processing and pattern recognition, an auxiliary constructor of scientific concepts and theories, and a simulator/replicator of the scientific discovery process. Simultaneously, it clarifies the epistemological boundaries of AI's involvement in scientific discovery from four dimensions: data dependency, causal reasoning, theoretical innovation, and psycho-social factors. Research indicates that while AI can expand the breadth and efficiency of scientific discovery, it cannot replace the core role of humans in theoretical innovation, value judgment, and complex causal modeling. Interdisciplinary integration is identified as the key pathway to transcend its current limitations.

**Keywords:** Artificial Intelligence; Scientific discovery; Epistemological boundaries

---

版权所有 © 2025 本文作者和香港科技出版集团。本作品根据知识共享署名国际许可证 (CC BY 4.0) 获得许可。 <http://creativecommons.org/licenses/by/4.0/>



**Open Access**