

何种科学变革将至

——从 AI 驱动的科学理论与理论科学的关系看

阮凯

【内容摘要】 人工智能驱动的科学研究的兴起与广泛应用，向科学哲学提出了一系列重要问题：AI 驱动的科学将带来哪些新变化？如何对其进行理论定位？它是否会引发新的科学范式变革？如果会，何种范式变革将至？要回答这些问题，首先需要厘清 AI 驱动的科学理论与理论科学的关系。经典科学哲学围绕科学的任务、界限、方法与发现逻辑等重大问题展开讨论，波普尔等科学哲学家强调科学由问题驱动，以描述和研究实在为任务，并充分阐述了观察负载理论、科学解释、因果性等议题。科学哲学理论与当代科学实践的对照研究表明，AI 驱动的科学理论与理论科学是一种融合互补的关系，它为科学研究提供了新工具和新方式。AI 驱动的科学目前虽然不能带来库恩意义上改变世界观的范式革命，但是从科学知识生产的方式、效率、类型乃至主体等细节看，它带来的变化也足以称得上是一种局部和弱化的科学范式变革。

【关键词】 AI 驱动的科学 理论科学 科学范式变革 波普尔哲学

【作者】 阮凯，中山大学哲学系（珠海）副教授。（广东珠海 519000）

【基金项目】 国家社科基金重大项目“数字智能技术与哲学发展及知识生产范式变革研究”（24&ZD320）

随着 AlphaFold 研究者德米斯·哈萨比斯（Demis Hassabis）和约翰·M·朱珀（John M. Jumper）被授予 2024 年诺贝尔化学奖，以表彰他们在蛋白结构预测方面的贡献，AI 驱动的科学（AI for Science）再一次受到学术界和社会的广泛关注，这也彰显了以大数据和机器学习为基础的人工智能技术与科学研究的深度融合发展，是人工智能和科学研究两大领域的前沿热点议题。2024 年 5 月，英国皇家科学院发布了重量级报告《人工智能时代的科学：人工智能如何改变科学研究



的性质和方法》，表明学术界已愈发重视 AI 驱动的科学带来的影响与变革。如何把握 AI 驱动的科学理论与理论科学的关系，思考人工智能是否以及如何改变科学的性质和方法，是哲学工作者要解答的时代之问。

一个对立的争论：我们如何定位人工智能驱动的科学

AI 驱动的科学是人工智能与科学的交叉融合，也是“人工智能+”最具前景的发展方向之一，这种融合使得“阿尔法折叠”（AlphaFold）以及利用机器学习来排除引力波噪声干扰等 AI 驱动的科学实例展现出变革性力量。有研究专门梳理了 AI 驱动的科学代表性案例，如“化学合成途径规划、生物医学序列的语言建模、超分辨率三维活细胞成像、粒子碰撞中的稀有事件筛选、核聚变反应堆的磁控制、在假设空间中导航等”，^①这些新进展大大提升了科学研究的效率、准确性和自动化程度，使我们有理由憧憬新一轮科技革命和科研范式变革。

但要回答何种科学变革将至，需要我们思考新兴科学实践如何改变传统科学实践。从传统科学发展到 AI 驱动的科学，一个重要变化是假说-理论驱动的科学已转换为数据-算法驱动的科学。理论常被哲学家认为是传统科学的核心，如恩格斯用非常肯定的语气说：“只要自然科学运用思维，它的发展形式就是假说。”^②20 世纪的物理学革命以来，这种看法尤为普遍，波普尔和库恩都发展了理论优位的科学哲学，两人把理论物理学视为经验科学的典范。波普尔写道：“现代理论物理学是直到目前为止我称作‘经验科学’的最完全的体现。”^③库恩则系统阐发范式革命，认为科学范式革命的核心部分往往是最具颠覆性的理论革命。在此背景下，AI 驱动的科学呈现很多新变化，它主要是指借助大数据和机器学习技术进行科学研究，其特点主要有：依赖科学数据、算法和算力，对大规模高质量数据的需求比对理论的需求更为显著。有学者也尝试概括其特点：“由数据驱动，具有很强的归纳性，并且相对独立于理论。”^④面对理论科学与 AI 驱动的科学之间大大小小的差异，当我们继续追问，AI 驱动的科学将带来何种科学变革？如何在整体上对其进行理论定位？我们发现，这些问题并没有现成答案，也容易形成相互对立、较为极端的学术立场。

过度鼓吹 AI 驱动的科学将带来彻底的范式革命的研究者主张替代论观点，即 AI 驱动的科学会取代理论科学，美国科技文化杂志《Wired》（《连线》）曾发表克里斯·安德森（Chris Anderson）的文章《理论的终结》，文章认为：“理论已结束：数据洪流使科学方法过时……当数据足够庞大，数字自会昭示一切。”^⑤这种替代论的相近表现形式就是数据崇拜，如图灵奖获得者朱迪亚·珀尔强调：“某些领域存在着一种对数据的近乎宗教性的信仰。这些领域的研究者坚信，只要我们在数据挖掘方面拥有足够多的智慧和技巧，我们就可以通过数据本身找到这些问题的答案。”^⑥而 AI 驱动的科学的一些严厉批评者支持非科学论，即 AI 驱动的科学并非真科学。科技哲学家沃尔夫冈·皮奇（Wolfgang Pietsch）很好地概括了这两种互相对立的观点：“大数据和数据科学中有一个认识论的信条，即所谓的‘相关性取代因果关系’。不仅大数据的支持者为这一论点辩护，而且大数据的批评者也常常利用它来展示大数据方法所谓的非科学性本质（unscientific nature）。”^⑦这句话提醒我们，针对人工智能在处理数据相关性方面的特长，不同的人会得出不同的结论，大数据的支持者倾向于认为大数据会在科学方面横扫一切；而大数据的批评者坚持认为，用大数据来做的研究并非科学研究。很明显，这两类主张构成了一个对立的争论。我们需要以学理的方式

① Hanchen Wang, Tianfan Fu, Yuanqi Du, et al., “Scientific Discovery in the Age of Artificial Intelligence,” *Nature*, vol. 620, 2023, p.48.

② 恩格斯：《自然辩证法》，北京：人民出版社，2018年，第110页。

③ 卡尔·波普尔：《科学发现的逻辑》，查汝强、邱仁宗、万木春译，杭州：中国美术学院出版社，2007年，第14页。

④ Wolfgang Pietsch, “Aspects of Theory-ladenness in Data-intensive Science,” *Philosophy of Science*, vol. 82, no.5, 2015, p.905.

⑤ Chris Anderson, “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete,” *Wired Magazine*, vol.16, no.7, 2008, pp.1-3.

⑥ 朱迪亚·珀尔、达纳·麦肯齐：《为什么：关于因果关系的新科学》，江生、于华译，北京：中信出版集团，2019年，第325页。

⑦ Wolfgang Pietsch, *Big Data*, Cambridge: Cambridge University Press, 2021, p.51.

回应这两个对立立场，但更重要的问题是：如何系统阐明 AI 驱动的科学理论与科学的关系，从而更准确地定位 AI 驱动的科学，并进一步揭示何种科学变革将至。

回应这些学术问题与争议，需要我们采取理论结合实践的研究方式，即一方面要考察 AI 参与科学实践的最新成果，从实践中提炼思想、解决争议；另一方面也要注重激活传统科学哲学思想资源，让经典科学哲学家“活在今天”，用他们的思想理论来审视科学研究的新变化。本文重点借鉴波普尔哲学，兼顾其他重要科学哲学家的相关思想，主要理由是波普尔对理论驱动的科学、科学发现的逻辑有系统的论述，他是理论优位的科学哲学的代表人物，他对科学的目标、科学的解释作用、实在论问题也有全面深刻的哲学阐发，有利于我们深入比较两类科学。

AI 驱动的科学理论与科学有何共性

要系统回答 AI 驱动的科学理论与科学有何关系，一个重要方向是思考它们是否有共性和联系，我们可以提出两个命题来尝试回答该问题，命题一是 AI 驱动的科学依然是问题驱动的科学；命题二是 AI 驱动的科学依然是描述和研究实在的科学。与此同时，对于两类科学共性与联系的讨论也有助于我们思考“人工智能驱动科学是非科学”论断的问题，因为“非科学论”往往站在理论科学的角度批评 AI 驱动的科学，否认两者的共性与联系。

（一）AI 驱动的科学依然是问题驱动的科学

按照波普尔的观点，传统的理论科学、经验科学都是问题驱动的科学。波普尔在《客观的知识》中提出：“我们不是从观察开始，而总是从问题开始，它们或者是实际问题，或者是已经陷于困境的理论。”^①波普尔的观点有其合理性，这体现在如下两点。

①③④⑤ 卡尔·波普尔：《客观的知识：一个进化论的研究》，舒炜光等译，杭州：中国美术学院出版社，2003年，第260页，第260—261页，第260页，第261页。

② 卡尔·波普尔：《科学发现的逻辑》，第xi页。

其一，问题是科学研究的逻辑起点，为科学研究提供方向和动力。从科学的主体看，科学是人类的科学，因此科学的核心目标也是解决人的问题，科学史研究一再表明，解决人类生存发展中的关键问题是科学发展的重要推动力。在波普尔眼中，所有哲学与科学问题中有一类问题尤为重要，也是哲学与科学魅力之所在，这就是宇宙论问题：“理解世界——包括作为世界一部分的我们自己和我们的知识——的问题。”^②从科学发现的逻辑看，问题的出现可以让人类进一步寻找解决问题的方法和路径，相比之下，无问题的观察容易使人陷入盲目。世界中有无数个对象和现象有待观察，理论上科学家也可以做无数观察，在科学研究中，我们凭什么说一个观察比另一个观察更为重要？虽然波普尔也看到观察的重要性，以及正是一些观察导致了问题的产生，但他仍坚持问题是科学研究的逻辑起点，他认为：“由于逻辑的理由，观察不能先于所有问题，虽然观察对于某些问题常常是明显居先的。”^③

其二，问题与猜想和反驳之间有更紧密的逻辑关系。观察不能直接引发猜想，只有观察引发了问题，我们才会提出某种猜想、假设和理论，以尝试去解决该问题，同时，因为猜想的出现，我们也会进一步对猜想进行反驳和批评。所以波普尔对科学知识的增长做了如下总结：“知识的增长是借助于猜想与反驳，从老问题到新问题的发展。”^④猜想与反驳的方法是“知识的增长所依据的方法”。^⑤人们为了解决问题，不断提出猜想与反驳，最后促进科学知识增长，构成了科学发展的一条逻辑主线。

像波普尔这样的经典科学哲学家将问题看作科学研究的逻辑起点，那么在 AI 驱动的科学中，情况是否也是如此呢？让我们首先考察一下 AI 驱动的科学代表性实例。2024 年 DeepMind 与

Isomorphic Labs 的研究人员共同在《自然》杂志上发表了关于 AlphaFold3 的论文, 文中指出: “该模型采用了大幅更新的基于扩散的架构 (diffusion-based architecture), 能够预测包括蛋白质、核酸、小分子、离子和修饰残基在内复合体的联合结构。”^①为何人类要用人工智能去预测这些生物分子结构及其相互作用呢? 这篇重磅论文正文的第一句话就揭示了原因: “生物复合体的精确模型对于我们理解细胞功能以及合理设计治疗方法至关重要。随着 AlphaFold 的发展, 蛋白质结构预测取得了巨大的进步, 许多后续方法也都建立在 AlphaFold 2 的理念和技术之上, 使该领域的发展突飞猛进。”^②如何理解细胞功能、如何为某个疑难杂症设计合理的治疗方法等都是重要的科学问题, 这些问题又进一步向科学家提出新的问题, 如何经济、效率并尽可能准确地掌握蛋白质等生物复合物的结构模型。因此, 从实践层面看, 困扰人类的科学问题依然是 AI 驱动的科学逻辑起点。

AI 驱动的科学是问题驱动的科学, 这与理论科学起源于问题、由问题驱动是一致的。换言之, 假说—理论驱动的科学和数据—算法驱动的科学都统一于问题驱动。因为科学的发展、建制的完善和共同体的形成, 各门科学都有自己的核心科学问题, 同时科学问题并不是孤立存在的, 很多科学问题往往交织关联, 这使得当前的 AI 驱动的科学不仅由问题驱动, 更是由科学问题驱动。很多问题在科学共同体中早已被公认是重要的科学问题, 随着人工智能技术的发展, 科学家可以利用人工智能更好地解决此类问题。要言之, 两类科学在科学问题上的承袭关系和连续性, 是为 AI 驱动的科学研究辩护的一个核心理由。需要强调的是, 在 AI 驱动的科学中, 不是人工智能自己提出问题再去解决问题, 而是人类利用人工智能技术去解决自己的科学问题。人类依然是提出科学问题的主体, 这在根本上保证了 AI 驱动的科学仍然是服务于人类、为人类谋福祉的事业。

(二) AI 驱动的科学依然是描述和研究实在的科学

在开端上科学主要由问题驱动, 但科学研究不会仅仅停留于提出问题, 为了更好地解释关于世界的谜题, 解决理论和实践问题, 科学家会提出描述事实世界和外部实在的陈述或理论, 问题驱动与描述实在共同构成了科学的两大主要特征。波普尔强调科学的两大特征是兼容的: “科学理论不仅仅是工具, 它同时也是真实的摹状陈述。它们是关于世界真实的猜想……这是实在论的立场。”^③实在论立场几乎贯穿波普尔整个科学哲学生涯, 在 1934 年出版的《科学发现的逻辑》中, 波普尔就说自己相信形而上学实在论, 在 1951 年至 1956 年写作的《实在论与科学的目标》中他更坚定了该立场, “我现在仍然相信形而上学实在论”, 其科学意义是: “它形成了某种背景, 赋予我们探索真理的行为以意义。旨在接近真理的理论讨论——也即批判性论证——如果没有客观实在性就毫无意义; 客观实在的世界正是我们探索的目标。”^④客观实在、真实世界和客观规律的存在, 是人类探索世界求取知识的前提, 如果不承认前者存在, 科学也就失去了研究的对象; 如果自然和社会不真实存在, 难道还会有自然科学和社会科学吗? 在 1972 年出版的《客观的知识: 一个进化论的研究》中, 波普尔继续主张描述实在是科学的任务, 他写道: “自然科学连同它们解决问题的批判方法, 以及一些社会科学, 特别是历史学和经济学, 相当长的时期以来表现着我们在解决问题和发现事实 (所谓发现事实, 我的意思当然是发现符合事实的陈述或理论) 方面的卓越努力。因此, 从真理的观点看, 这些科学总的来说包括最优越的陈述和理论: 即包括对事实世界或所谓的‘实在’提供最好描述的那些陈述或理论。”^⑤

波普尔从实在论立场出发进一步概括了科学的目标与任务, 一些支持实在论的哲学家也通过反驳反实在论进一步补充了波普尔的观点, 如 A. F. 查尔默斯认为: “尽管确实, 我们不使

① ② Josh Abramson, Jonas Adler, Jack Dunger, et al., “Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3,” *Nature*, vol.630, 2024, p.493.

③④ 卡尔·波普尔:《实在论与科学的目标》, 刘国柱译, 万木春校, 杭州: 中国美术学院出版社, 2008 年, 第 105 页, 第 74—75 页。

⑤ 卡尔·波普尔:《客观的知识: 一个进化论的研究》, 第 288 页。



① A. F. 查尔默斯：《科学究竟是什么》，鲁旭东译，北京：商务印书馆，2018年，第262页。

② Kaifeng Bi, Lingxi Xie, Hengheng Zhang, et al., "Accurate Medium-range Global Weather Forecasting with 3D Neural Networks," *Nature*, 2023, vol.619, p.533.

③ 卡尔·波普尔：《客观的知识：一个进化论的研究》，第288页。

用某种概念框架就无法描述世界，但我们仍然可以通过与世界的相互作用来检验那些描述的适当性。”^①由此，科学哲学家的相关观点与论证启发我们，从实在论的角度为AI驱动的科学辩护。

第一，AI驱动的科学是描述和认识实在的科学。波普尔一直强调科学需要研究客观实在和真实世界，他在《科学发现的逻辑》中主要强调的是科学理论对于实在的描述，常把科学知识 with 科学理论联系在一起。他如此重视理论与其反对归纳方法和提倡假设—演绎法密切相关，理论是全称陈述，虽不能由单称陈述归纳而来，但理论或假设可被经验检验和证伪。在《客观的知识》中，他更严谨地说科学是包括了最优越的“陈述和理论”，陈述或理论是对于实在的描述。这里的“陈述”自然包括了单称陈述和存在陈述。

波普尔对科学阐释的微妙变化启示我们，虽然科学理论是人类描述和认识实在的重要形式之一，但如果我们把理论作为唯一形式，就过于武断了，描述和认识实在有不同的形式。比如，事件视界望远镜（EHT）合作组织发布的首张黑洞照片也是描述和认识实在的形式。就AI驱动的科学而言，虽然由算法生产出来的预测模型（如预测生物大分子的结构模型）、分类结果（如对引力波信号与噪声进行分类）等不具有理论形态，但它们仍然属于人类描述和认识实在的形态和方式。从AlphaFold到AlphaFold 3的一系列研究成果都是为了愈发准确和广泛地预测蛋白质及更多的生物分子的结构及其相互作用，这本质上仍然体现了科学发现事实、描述实在的工作。AI驱动的科学也能帮助人类更好地描述客观实在、认识未知世界。

第二，AI驱动的科学能较好地结合解决科学问题与描述客观实在。波普尔之所以把社会科学和自然科学放在一起讨论，原因是两者既包含解决问题的维度，同时也有发现事实、描述实在的维度。一般而言，科学由问题驱动与科学描述客观实在两者是兼容的，像物理学和化学等基础科学不仅很好地描述了外部世界，同时这些对自然的卓越描述也很好地服务于人类在解释自然、利用自然中遇到的新问题。在时间维度上，科学对世界的描述并不局限于当下，如考古学是对过去世界的描述，天文学、气象学也包括对世界未来状态的描述和预测。以AI驱动的气象科学为例，如何准确预测天气既是一个重要的科学问题，同时也表现为对世界的描述和预测。我国研究人员2023年开发了“盘古天气”（Pangu-Weather），开发者指出，“与世界上最好的数值天气预报（NWP）系统——欧洲中期天气预报中心（ECMWF）——的集成预报相比，盘古天气在所有测试变量的再分析数据上取得了更强的确定性预报结果”。^②这进一步表明，AI驱动的气象科学在解决问题和描述实在、预测事实方面与传统气象科学具有连续性，新技术运用导致了预测精度的进一步提升，这主要是一种程度层面的变化。

在科学哲学理论与具体实践案例的结合中，我们可以进一步从实在论层面为AI驱动的科学辩护：AI驱动的科学是人类科学大厦的一部分，它仍然延续了传统科学“描述实在、认识实在、揭示实在”的追求，呼应了科学哲学中的实在论传统。例如，AlphaFold预测和描述的不是某个虚构的蛋白质模型，而是真实存在的蛋白质模型；盘古天气预测和描述的也不是电子游戏中的虚拟天气，而是现实世界未来几天内真实的天气情况。波普尔也在《客观的知识》中将科学对实在、真实、事实的追求视作科学的理想，他认为，科学一直受到一个理想的鼓舞，即“寻求科学问题的真答案：符合事实的答案”。^③总之，尽管AI驱动的科学呈现新特征，但它仍以解决问题并认识实在为基本任务，在研究目标上与理论科学保持着深刻的连续性，质疑AI驱动的科学本质上是非科学，这一观点是站不住脚的。

人工智能驱动的科学能否取代理论科学

问题驱动和描述实在这两个代表性特征已显示两类科学的延续性，正是因为 AI 驱动的科学保留了传统科学的一些核心特征，我们才将其视作科学。在回应非科学论的同时，我们还要应对替代论的挑战，进而合理地评价这一新兴科学形态，全面总结两类科学的关系。AI 驱动的科学究竟能不能取代理论科学？这并不是一个简单的是或否的问题，我们更需要追问能否取代的充分理由。因此，我们需要考察 AI 驱动的科学带来的新变化，深入研究这些新变化有助于揭示它的潜力和局限，也有助于进一步审视它与理论科学的关系。

AI 驱动的科学具备传统科学不具备的一些优点，这包括但不限于：一定程度的自动化，它能够实现一定程度的自动化预测和科学发现；擅长利用大数据做出相关性分析。它与统计学的亲缘关系使得 AI 驱动的科学擅长从大数据中发现相关性信息，进而使自动化的科学分类和预测成为可能。各种优质科学数据库和算力中心的建立也为未来更广泛的 AI 驱动的科学研究提供了物质基础。虽然 AI 驱动的科学具备如上优势，使其在某些方面比理论科学做得更好，更能满足科学家的研究需求，但 AI 驱动的科学能真正取代理论科学进而开辟一个科学新纪元吗？笔者希望通过如下三方面的论述来反驳替代论的观点。

（一）从科学发现负载理论驳斥替代论

波普尔清楚地认识到，在科学发现的诸环节，理论都发挥着重要作用。从经验观察到记录数据再到验证理论，这些环节都不同程度地负载理论。就观察负载理论和语言负载理论而言，他明确反对排除理论观察和清空理论的现象学语言。他在《科学发现的逻辑》的一处易被忽略的脚注中写道：“我们的日常语言是充满着理论的，观察总是借助于理论观察；只有归纳主义者的偏见才使得人们认为：可能有一种现象语言，不包含理论，可以和‘理论语言’区别开来。”^①就验证负载理论而言，他主张：“即使实验家，他的大部分工作也不是进行精确的观察，他的工作也主要是理论性的。理论支配着实验工作，从它开始计划一直到在实验室里最后完成。”^②理论是实验的前提与基础，没有理论，实验就找不到方向。有了理论和假设，我们才会找到检验理论的方向。

在波普尔之后，越来越多的科学哲学家也从不同角度论证了观察与数据负载理论的观点。例如，凭借《以数据为中心的生物学：一个哲学研究》(*Data-Centric Biology: A Philosophical Study*) 获得 2018 年拉卡托斯奖的萨比娜·莱奥内丽 (Sabina Leonelli) 在该书中探讨了数据的理论负载问题：“数据是人造的产物。数据是研究者与世界之间复杂交互过程的结果，这一过程借助于观测技术、记录和测量设备等界面，对研究对象（如有机样本甚至整个生物体）进行重新标度、修改和标准化，使其适合进行研究。”^③莱奥内丽也在论证数据负载理论时引用了以提出观察渗透理论而闻名的美国哲学家诺伍德·汉森 (Norwood R. Hanson) 的名言：“看的行为就是负载理论的工作 (theory-laden undertaking).”^④从波普尔到汉森再到莱奥内丽，一系列科学哲学家的努力让我们看到，观察负载理论与数据负载理论是一脉相承的。

科学哲学对观察负载理论、数据负载理论的强调也得到了 AI 驱动的科学实践的印证。首先，具体的 AI 驱动的科学研究究竟需要什么类型的数据，离不开科学家的理论观点。正如理论科学不能从随机观察开始一样，大数据科学也不能从随机数据开始，是科学家基于科学理论、科学知识来决定用哪些数据。比如，AI 驱动的气象研究需要的是专业的气象数据，而不是生物、天文

①② 卡尔·波普尔：《科学发现的逻辑》，第 35 页，第 84 页。

③④ Sabina Leonelli, *Data-Centric Biology: A Philosophical Study*, Chicago: University of Chicago Press, 2016, pp.71-72, p.72.

①② John Jumper, Richard Evans, Alexander Pritzl, et al., "Highly Accurate Protein Structure Prediction with AlphaFold," *Nature*, vol. 596, no. 7873, 2021, p.587, p.584.

③④ 卡尔·波普尔:《科学发现的逻辑》,第14页,第253页。

⑤ 卡尔·波普尔:《客观的知识:一个进化论的研究》,第42页。

⑥⑦ Thomas S. Kuhn, *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press, 1962, p.55, p.55.

等其他领域的任何数据。其次,专业科学数据的生产也是以科学理论为前提的。按照研发者的介绍,“AlphaFold 架构通过使用 PDB 中的数据进行监督学习来提升预测精度”,^①但蛋白质结构数据库(PDB)中的数据基本都是结构生物学家通过 X 射线晶体学、冷冻电子显微镜等科学技术测定出来的,而这些科学技术的核心都是科学理论。最后,数据负载理论会影响算法负载理论。因为大数据科学的核心是数据和算法,所以数据负载理论会使得算法也负载理论。具体而言,一方面,科学家用数据来设计和验证算法模型,使算法间接负载理论;另一方面,科学家在设计算法时也需要利用科学理论来提升算法的表现,使算法直接负载理论。如在 AlphaFold 2 算法设计中,研究者为了使算法具备更优秀的预测能力,将人类关于蛋白质的理论知识融入算法设计中:“AlphaFold 以蛋白质结构的进化的、物理的和几何的约束条件为基础,开发设计新型神经网络架构与训练程序,大大提高了对蛋白质结构预测的准确性。”^②背景性的科学理论使科学家在设计算法时得到了更多科学指导,算法的表现也因此获得巨大提升。

由此可见,虽然理论科学与 AI 驱动的科学在运作模式上有所不同,但后者也并非运行在理论的真空中,所以用 AI 驱动的科学来完全取代理论科学是行不通的;否则,它在抛弃理论的同时,也把自己一起抛弃了。

(二) 从解释的局限驳斥替代论

从理论科学到 AI 驱动的科学的变化,直接导致了科学研究从解释充分到解释不足的变化。一般而言,科学理论的一个主要功能是科学解释,发挥科学理论的解释性是传统科学的一个主要特点。比如,爱因斯坦的狭义相对论解释了时间和空间的相对性,广义相对论对引力做了几何解释。我们为了在世界中更好地生存和发展,需要尽可能地去认识和理解我们所处的世界,而科学理论既是我们认识世界的结晶,也是我们解释世界的有力工具。波普尔用一个形象的比喻概括了科学理论的解释作用以及为何人类非常需要科学:“理论是我们撒出去抓住‘世界’的网。理论使得世界合理化,说明它,并且支配它。我们尽力使这个网的网眼越来越小。”^③由于人的思想、假设和猜想都与理论密切相关,所以波普尔也尤为重视它们的解释作用,他写道:“大胆的想法,未被证明的预感,以及思辨的思想是我们解释自然的唯一手段。”^④

人类借助科学理论去解释我们所处的世界,而且随着科学的发展,解释会愈加缜密精细,解释的“网眼”越来越小。比如,达尔文用自然选择理论解释生物的演化问题,但随后 DNA 的发现为人类理解遗传变异提供了更精细的科学解释,1953 年沃森和克里克提出的 DNA 双螺旋结构理论进一步解释了遗传信息如何进行复制、传递与变异。波普尔十分重视科学的解释作用,以至于将解释实在与描述实在并列,都视为科学的主要目标。他认为:“在科学中我们所力图做到的是描述和(尽可能地)解释实在。我们借助猜测性理论达到这一点……”^⑤波普尔之后,库恩也强调利用概念与理论框架去解释科学现象。他写道:“发现一种新的现象本质上是一个复杂的过程,它既包括认识到某个东西存在,也包括认识到它是什么(what it is).”^⑥在理论科学中,我们并不满足于知道某个科学研究对象(如黑洞、引力波)存在或不存在这样简单的科学答案,更需要掌握科学答案背后更具体的科学解释,比如现象背后的因果机制是什么,如何在既有的科学理论框架下理解新现象与新实体,这些都是“是什么”的问题。所以库恩紧接着指出:“观察和概念化、事实和将事实同化到理论中,在发现过程中是不可分割地联系在一起。”^⑦这些观点反映了经典科学哲学家都非常重视科学的解释作用,他们认为科学解释离不开人类的概念系统和理论系统。

科学史的发展一再证明,理论科学具有非常强大的解释力,相比之下,在 AI 驱动的科学实

践中,解释虽然存在,但远远不够充分,而且为数不多的解释也是科学家基于科学理论所作的解释。这使得 AI 驱动的科学替代理论科学的观点自相矛盾。通常来说,算法和机器的计算过程对人来说是认知黑箱,如董春雨教授指出,“多层级神经网络计算对认识不透明性的限制是难以克服的。”^①而机器认识的不透明性与 AI 驱动的科学实践中算法难以提供有效科学解释是紧密相关的,试图用不透明的算法来清楚地解释世界的运行规律、新的科学现象和科学实体,这几乎是不可能完成的任务。当然,如果我们审视 AI 驱动的科学研究的全过程,会发现其实存在一定的解释,但它们大多是科学家做出的解释,发表在《自然》等学术期刊上有关 AI 驱动的科学论文,都可以理解为科学家向科学共同体和读者给出的全方位解释。在问题界定环节,研究者要解释他们研究的是什么领域的问题、该问题有何价值、该研究可以细化为哪些具体问题等;在算法设计环节,研究者要解释他们如何设计算法,以及在设计算法时融入了哪些科学理论;在算法评估环节,研究者要利用已有的数据等去重新评估算法的准确性,解释其应用范围和局限性等。

通过上述讨论我们发现,经典科学哲学家把解释世界作为科学的主要目标,而 AI 驱动的科学在科学解释上有天然的局限。这使得利用算法作出的预测无法取代理论知识的解释。同时,科学家在 AI 驱动的科学中开展的大量解释性工作,也以他们掌握的科学理论为前提。这两方面的原因使我们有充分的理由来驳斥替代论。

(三) 从因果性研究与相关性研究的互补来驳斥替代论

从科学发展史看,理论科学较为擅长因果性研究。从学理上看,理论为何擅长因果性研究呢?波普尔对此给出了一个实在论的解释,自然有“真正的规律性”,^②自然中真实存在的普遍性定律是因果解释的基础。因为规律具有普遍性,所以“科学理论是全称陈述”。^③相比之下,AI 驱动的科学并不需要预设理论科学研究的形而上学前提——自然有真正的规律性,机器学习与统计学之间存在紧密联系,^④使得 AI 驱动的科学更擅长进行相关性研究。相关性指的是两个或多个变量之间的统计关联,即使两个变量之间存在相关性,也不一定意味着一个变量是导致另一个变量变化的原因。

相关性分析虽然可以帮助科学家发现潜在的因果关系和科学规律,但它本身并不提供关于因果机制的直接证据,更不能替代因果性研究。朱迪亚·珀尔对此也有明确的判断:“借助贝叶斯网络,我们教会了机器在灰色地带进行思考,这是机器迈向强人工智能的重要一步。但就目前而言,我们仍然无法教会机器理解事情的前因后果。”^⑤他用一个生动的例子来表达自己的立场:“一个国家的人均巧克力消费量和该国诺贝尔奖得主的人数之间存在强相关。这种相关性显然是很愚蠢的,因为不管我们怎么想象,吃巧克力看起来都不可能导致我们获得诺贝尔奖。”^⑥一个相关关系远远不够,我们有时还需要海量的因果解释和理论说明,相关性研究的局限性由此可管中窥豹。因果性研究有不可被替代的重要理由:“在理性场域,即使否定因果性也必须依赖因果观念……理论不仅必不可少,没有理论,数据毫无价值,甚至没有意义。”^⑦因果性在科学中的不可替代性体现在多个方面。第一,因果性是科学理论的基础,很多科学理论都建立在因果性之上,无因果就无理论,相关性研究难以替代因果性研究来为这些理论奠基。第二,因果性知识有无可取代的实践效用,它能让人进行更精确的预测、控制和干预,例如在实验中更好地控制变量。第三,因果性知识是科学解释的基础,AI 驱动的科学在科学解释上的局限与其不善于因果性研究紧密相关。从研究实践看,研究人员已经认识到 AI 驱动的科学只依赖相关性分析的局限性,并开始开发新的方法和技术来增强人工智能在因果性研究方面的能力,比如一些研究者为了让机器提供人类能够

① 董春雨:《从机器认识的不透明性看人工智能的本质及其限度》,《中国社会科学》2023年第5期。

②③ 波普尔:《科学发现的逻辑》,第255页,第35页。

④ 这种联系也体现在它们共享的一些方法和核心概念上,如线性回归、贝叶斯方法、决策树、随机森林、聚类分析、置信区间等。

⑤⑥ 朱迪亚·珀尔:《为什么:关于因果关系的新科学》,第323页,第48页。

⑦ 王天思:《大数据中的因果关系及其哲学内涵》,《中国社会科学》2016年第5期。



① 参见 Yu-Liang Chou, Catarina Moreira, Peter Bruza, et al., "Counterfactuals and Causability in Explainable Artificial Intelligence: The Theory Algorithms, and Applications," *Information Fusion*, vol.81, 2022, pp.59-83.

② Wolfgang Pietsch, *Big Data*, p.69.

③ 刘大椿、赵俊海：《科学哲学的经验主义新建构》，《中国社会科学》2016年第8期。

④ 伊恩·哈金：《驯服偶然》，刘钢译，北京：商务印书馆，2015年，第3—5页。

⑤ 在气象学中，影响未来气候的因素实在太多，科学家要寻找一种预测性的理论非常困难，因而转向用更为复杂的计算模型来预测天气。

⑥ 参见 Sabina Leonelli, "What Difference Does Quantity Make? On the Epistemology of Big Data in Biology," *Big Data & Society*, vol.1, no.1, 2014, p.8.

理解的因果解释，使机器模拟人的反事实推理，用反事实算法 (counterfactual algorithms)^①赋予机器因果推理和因果解释能力，但这些研究尚不成熟，还没有代表性的 AI 驱动的科学研究成果问世。

上述分析都支持一种更为中道的观点：因果性研究和相关性研究在科学探索中应互相补充，而非彼此取代。原因是：第一，两种研究都在科学中取得成功，但每一项研究都不是科学事业的全部。沃尔夫冈·皮奇对大数据的研究，也以另一种表述佐证了互相补充立场的合理性。他指出：“事实上，数据科学和机器学习中的成功科学实践，构成了支持归纳主义认识论复兴的迄今为止最有说服力的论据……当然，这并不是说归纳主义的大数据方法对于回答所有科学问题都有用。相反，它们应该被看作是对现有科学方法工具箱的补充。”^②我们也可以从与皮奇相反的视角为补充论辩护：人类在理论科学方面的科学实践历经千百年的发展，已构建了令人依赖的科学大厦，理论取得的非凡成功也构成了支持因果性研究和假设-演绎方法的最有说服力的论据，而 AI 驱动的科学实践也表明，它更多是在理论基础薄弱和理论难以满足科研需求的领域，发挥了对理论科学的有效补充作用。1980 年代以来，科学哲学总体发展趋势是从理论优位的科学哲学走向实践优位的科学哲学，^③其内在逻辑是：理论虽然重要但并非科学的全部。从这种实践视角看，因果性研究与相关性研究，理论科学与 AI 驱动的科学，都能统一于人类实践中来，实践优位的科学哲学能将理论优位和数据优位的科学哲学包容进来。

第二，在一个科学的世界观中，必然与偶然并存，规律与随机并存，因此科学需要不同的研究方式。一方面，在成熟的理论科学中，科学家已经揭示了世界中存在的许多重要自然规律，并不需要算法模型去重复理论科学的工作，这对科学家而言，既无必要也不可行，比如目前没有科学家用算法模型来完成麦克斯韦方程组的工作。另一方面，世界中也存在排斥自然定律的偶然性和随机性，科学哲学家伊恩·哈金指出：“量子力学认为自然界在其基层具有无法克服的随机性……偶然定律颠覆了决定论。”^④在有些领域，用数学公式表达的自然规律很难发现，甚至规律压根不存在。例如，在气象科学等领域，定律性的理论较少，影响预测结果的变量和因素过多，导致科学家难以建立统一的气象定律和预测理论。^⑤科学家也没有提出科学理论来帮助人类从氨基酸序列自动推断蛋白质和生物大分子的结构，这时 AlphaFold 就可以起到补充的作用，为人类快速预测生物大分子结构提供帮助。现实中，AI 驱动的科学研究的的确是在理论科学做得不够的地方辅助人类认识世界，解决科学问题。

总体而言，AI 驱动的科学有其突出的优势，但我们也不能忽略其对理论的依赖、解释不足、不擅长因果研究、对高质量数据库的依赖等局限性。除了上文讨论的一些局限性外，我们也需要看到 AI 驱动的科学可能导致偏见，如萨比娜·莱奥内利认为，宣传大数据科学是全能且无偏见的观点，将误导科学家与公众。^⑥因篇幅原因，本文不再详述偏见等其他局限性。

AI 驱动的科学理论与理论科学的融合互补及其启示

AI 驱动的科学作为一个新生事物，为传统科学的发展带来了新的机遇和挑战。在波普尔科学哲学思想和前沿实践的双重重视域下，我们可以更好地比较分析 AI 驱动的科学与传统科学的异同，对科学研究的新形态进行理论定位。一方面，AI 驱动的科学是人类科学事业的一部分，在解决科学问题与描述实在上，它与理论科学共享着科学的基本精神和特征，并不是打着科学旗号的伪科

学或非科学；另一方面，AI驱动的科学不会替代理论科学，两者是融合互补的关系。互补关系前文做了总结，理论科学的内核是预设世界存在自然规律并研究自然规律，AI驱动的科学内核是在统计学基础上对未知世界进行预测。而融合关系体现在：AI驱动的科学需要理论科学的支持，需要将理论工作融入其中；同时AI驱动的科学也会参与和促进理论科学的发展，比如“借助AI提出科学假说”，^①当然目前这仍然是假说筛选、搜索和优化等初步工作。未来的科学研究将形成以理论科学为基础，理论科学与AI驱动的科学融合互补、互相促进、各显其长的科学研究新格局，这种新格局对思考科学哲学和展望科学未来发展有五点重要启示。

第一，理论科学与AI驱动的科学、假设-演绎法与归纳法、人类创造性思考与机器程序性计算都能对科学进步、知识增长做出贡献。在理论科学时代，提出创新性的科学假设与理论往往需要天才般的智力和灵感，所以波普尔既推崇假设-演绎法，也重视理论研究的创造性、猜测性，他认为：科学假设和科学理论的提出都需要非凡的创造力，“每一个科学发现都包含‘非理性因素’，或者在柏格森意义上的‘创造性直觉’”。^②无论我们是使用演绎逻辑还是归纳逻辑都无法从伽利略和开普勒理论中推导出牛顿力学，“只有天才般的独创性才能迈出这一步”。^③但AI驱动的科学，并不会对科学家的创造力和非凡的直觉有如此高的要求，它更多地依靠大数据与机器学习，是一种归纳方法的体现。沃尔夫冈·皮奇概括了两种科学方法与算法的关系：“归纳方法往往非常简单，能够以准算法（quasi-algorithmically）的方式执行，而假设的制定则需要创造力和直觉，这些是难以轻易地用算法来实现的。”^④两种科学的融合互补使得我们在科学方法论问题上更加开放包容，不必像一些科学哲学家那样，过于推崇某种科学方法而排斥其他的科学方法，假设-演绎法与归纳法都能以不同的方式推动科学的进步和科学知识的增长。这使得我们反过来又超越了波普尔哲学，克服其哲学的局限性。

第二，理论科学与AI驱动的科学各有所长，科学家可以根据研究条件和研究目标来选择适合的研究方式。基于研究条件来确定研究方式，是因为数据质量等研究条件会影响我们选择何种研究方式。从数据的量看，并非所有科学研究都有完善的数据库可供使用，且储存科学数据的大数据库并非唾手可得。比如，只有在人类建立了像蛋白质结构数据库后，用人工智能预测蛋白质三维结构才得以可能，而这些数据库又凝结了科学家艰辛的理论工作和海量的实验工作。我们难道要为每一种科学研究建立这样的数据库吗？这是不切实际的要求。从数据的质看，大数据并不意味着高质量数据，数据在生产、储存、传播、使用过程中会出现各种各样的问题，如错误数据被生产出来，数据储存时数据格式被改变，数据传播过程中关键数据丢失，数据因为缺乏解释而被误用等。因此，高质量的大数据既是AI驱动的科学取得成功的前提，也是限制其发展的原因之一。相比之下，理论科学家在解决科学问题时更少地受到数据的限制。比如爱因斯坦在提出相对论时，并没有掌握也不需要太多物理数据。历史上的很多科学家从小数据中提炼出伟大的猜想，提出经受住各种检验的科学理论，更显示了这些科学家的惊人智慧。

基于研究目标来确定研究方式，可以让不同科学研究方式扬长避短。AI驱动的科学能完成的很多工作，理论科学与实验科学往往能做得更好，反之亦然。如传统实验方式早已能测定蛋白质的三维结构，而AlphaFold只是通过算法预测蛋白质的三维结构，由于预测不能完全保证百分之百的准确性，所以预测的模型都会用置信度pLDDT（predicted Local Distance Difference Test）等方式来表示预测的准确性和可信度。为何实验方式更准确，我们还需要AI驱动的科学去预测？其实，AI驱动的科学之所以有发挥的舞台，是因为科学家有准确性之外的其他研究目标，比如在适

① Hanchen Wang, Tianfan Fu, Yuanqi Du, et al., “Scientific Discovery in the Age of Artificial Intelligence,” *Nature*, vol. 620, 2023, <https://doi.org/10.1038/s41586-023-06221-2>, p.51.

② 卡尔·波普尔：《科学发现的逻辑》，第8页。

③ 卡尔·波普尔：《实在论与科学的目标》，第138页。

④ Wolfgang Pietsch, *Big Data*, p.16.

当牺牲准确性的前提下提高效率 and 自动化程度。实验方式固然准确,但颇为费时费力,效率并不高,而 AI 驱动的科学在自动化程度、经济效率等方面有其优势,因此当科学家们需要经济、效率地掌握海量的蛋白质三维结构,同时恰好有适合的数据库等研究条件时,AlphaFold 才作为一种 AI 驱动的科学代表横空出世,有效地补充传统的实验模式。

第三,科学家也依然是 AI 驱动的科学研究的主体。理论科学需要科学家的理性与创造力,波普尔也强调“科学家的工作是提出和检验理论”^①,因此他们毫无疑问是理论科学研究的主体。但是 AI 驱动的科学自动化特征,很容易让我们忽视科学家的主体作用。在比较两类科学的异同之后,我们更有理由认为科学家是 AI 驱动的科学主体:一方面,AI 驱动的科学解决的是科学家和全人类的科学问题,而不是困扰人工智能的科学问题;另一方面,在 AI 驱动的科学各个环节中,科学家都起到至关重要的作用,科学家主要围绕数据与算法模型展开工作。这些工作包括:创建数据集、数据预处理、测试和验证模型、选择模型架构^②以及解释模型的预测结果。因此,数据和算法不是凭空产生的,而是科学家集体劳动的成果,AI 驱动的科学不是机器的事业,而是人类的事业。

第四,AI 驱动的科学研究实践是我们理解科学,促进科学哲学未来发展的重要推动力。虽然我们看到波普尔哲学对于审视 AI 驱动的科学有一定的理论指导意义,且新的科学形态背后的科学哲学方法和立场与理论科学并不矛盾;但是我们也要意识到,科学哲学是对科学实践的理论总结与哲学反思,其自身也是在不断发展的。随着 AI 驱动的科学等新型科学实践的发展,传统的科学哲学理论日益显露出其不充分性。这启示我们:一方面要不断根据科学发展与时俱进地更新对于科学的理解,提出更全面深刻、经得起实践检验的科学哲学理论。例如,波普尔将科学理论当作科学知识的主要形态,就不能完全说明 AI 驱动的科学研究产生的知识,如科学理论、科学事实、科学家利用机器算法做出较为准确的科学预测与分类都可被视作科学知识,它们只是科学知识的不同类型而已。另一方面,不要将经典的科学哲学观点简单地套用在新兴科学实践上。例如,虽然波普尔建议将“可证伪性(falsifiability)作为我们判定一个理论系统是否属于经验科学的标准”,^③但我们不能简单地拿可证伪性标准来衡量 AI 驱动的科学研究。在大数据科学中,虽然测试与验证算法和验证理论有相似性,但是科学家并不会因为算法给出了错误或不准的预测而放弃算法,而是不断地改进算法以提升预测精度。

第五,AI 驱动的科学带来的是拓展科学研究方式、丰富科学知识形态、增强自动化水平的局部科学变革,而非库恩意义上的“改变世界观”(changes of world view)^④的科学革命。我们既不能忽视这项影响深远的科学变革,因为科学家的确开拓了认识自然、改造世界、解决问题的新工具和新方式;也不能过分夸大这种科学变革的范围和影响,以至于主张 AI 驱动的科学替代论和理论消亡论。理论科学与 AI 驱动的科学将构成双向赋智、互相促进的关系,前者为后者的数据预处理、算法设计、预测结果解释等提供理论支持,后者帮助前者更高效地分析处理大规模数据、提出理论假设、模拟实验过程、理解理论的局限性等。两者共同推动人类的科学事业迈上一个崭新的台阶。库恩对范式不可通约性的强调,容易引导人们以为新科学变革要与以往科学切断联系,如果我们把范式理解为一整套研究方式,两类科学双向赋智、互相促进的新科学何尝不是一种新的科学范式。

编辑 张 蕾

①③ 卡尔·波普尔:《科学发现的逻辑》,第7页,第54页。

② Jens Ulrik Hansen, Paula Quinon, “The Importance of Expert Knowledge in Big Data and Machine Learning,” *Synthese*, vol.201, no.35, 2023, pp. 10-14.

④ Thomas Kuhn, *The Structure of Scientific Revolutions*, p.111.