

科研智能：政策、技术及应用

董昊, 魏凯

中国信息通信研究院, 北京 100083

摘要

如今, 科研智能已成为一个高度交叉融合的领域, 其整合了多个学科的知识与方法, 广泛应用于科学研究和产业研发。深入探讨科研智能的起源、发展及未来趋势: 首先, 界定科研智能的概念, 明确其范畴与核心特征; 其次, 分析全球主要经济体在政策层面的战略布局; 接着, 介绍科研智能的关键技术及其在不同科研领域的实践成效; 最后, 展望科研智能未来的演进方向。科研智能将深刻改变科学发现模式、重塑技术创新流程, 并为解决人类社会面临的重大挑战提供重要支持。

关键词

AI; 科研智能; AI4RD; 科学智能; AI4S

中图分类号: TP18

文献标志码: A

doi: 10.11959/j. issn. 2096-0271.2025084

AI4RD: policy, technology, and applications

Dong Hao, Wei Kai

China Academy of Information and Communications Technology, Beijing 100083, China

Abstract

Artificial intelligence for research and development (AI4RD) has evolved into a highly interdisciplinary field, integrating knowledge and methods from diverse disciplines with broad applications in scientific research and industrial development. This paper will delve into the origins, development, and future trends of AI4RD. First, we will define the concept of AI4RD, clarifying its scope and core characteristics. Next, we will analyze the strategic policies of major global economies in this area. Following that, we will introduce the key technologies of AI4RD and examine its practical outcomes in different research fields. Finally, we will explore the future evolution of AI4RD. We believe that AI4RD will profoundly transform scientific discovery methods, reshape technological innovation processes, and provide significant support in addressing major challenges faced by human society.

Key words

AI, artificial intelligence for research and development, AI4RD, artificial intelligence for science, AI4S

0 引言

科 研 智 能 (AI for research and development, AI4RD) 作为人工智能与科学研究及产业研发交叉融合的前沿领域,近年来迅速崛起,成为推动科学发现和技术创新的重要动力。科研智能概念的起源,可追溯至计算机科学的早期,当时科研人员开始尝试利用计算机进行数值计算、数据分析和模拟仿真,以辅助传统的实验研究和理论研究。随着计算能力的飞速提升和算法的不断演进,计算范式逐渐成为与实验范式、理论范式并列的“第三种科研范式”。进入21世纪,互联网的普及和数据的爆炸式增长催生了“电子科学(e-science)”和“数据密集型科学(data-intensive science)”等理念,计算机科学家吉姆·格雷(Jim Gray)将其概括为“第四种科研范式”——数据驱动科研范式,强调通过大数据分析 with 挖掘推动科学发现^[1]。

科研智能的演进历程呈现出明显的阶段性特征:第一阶段(1950—1999年)以数值计算和模拟仿真为核心,奠定计算科学基础;第二阶段(2000—2010年)依托数据驱动方法,推动第四范式形成;第三阶段(2010年至今)在深度学习革命推动下,实现智能方法与科研实践的深度融合。

真正意义上的科研智能是近十年来伴随人工智能技术的突破性进展——特别是深度学习和大语言模型(large language model, LLM, 简称大模型)的兴起——而蓬勃发展起来的。早在2018年,鄂维南院士提出“AI4S (AI for science)”的概念,并将其译为“科学智能”,即利用人工智能解决科研中的复杂问题,加速科学发现。2020年,DeepMind 发布的 AlphaFold2 首次实现了

接近实验精度的蛋白质结构预测,成为科研智能领域的里程碑式成果^[2]。同年,美国能源部发布《科学智能研究报告》,系统梳理人工智能在科学研究中的关键议题^[3]。李国杰院士团队提出“第五种科研范式”的概念,并在后续研究中将其定义为“智能化科研(AI for research)”,进一步明确了人工智能在基础科学、技术研究和工程应用中的赋能价值^[4-5]。自2021年起,北京科学智能研究院、法国国家科学研究中心人工智能科学与科学人工智能中心、微软研究院科学智能研究中心、康奈尔大学科学智能研究所、上海科学智能研究院等机构相继成立,专注于人工智能与科研的融合创新。斯坦福大学、麻省理工学院、中国科学院、清华大学、北京航空航天大学等知名高校和研究机构积极布局,谷歌、英伟达、微软、华为、百度、阿里等科技企业纷纷加入, Atomwise、Insilico Medicine、北京深势科技有限公司、深圳晶泰科技有限公司(简称晶泰科技)等创新企业快速崛起,相关机构和企业 在 2022 年至 2023 年密集发布了 AI+生物医药、AI+数学、AI+材料、AI+气象、AI+核聚变等领域的研究成果。自2023年起,中国、美国和欧盟等主要经济体分别从政策、法规及资金支持等维度体系化推动科研智能发展,行业组织也开始积极布局,如中国人工智能产业发展联盟成立科学智能工作组。2024年10月,诺贝尔奖首次颁发给科研智能成果(AlphaFold系列模型),象征着人工智能在科研中的地位正式迈向主流,预示着由人工智能驱动的科研新时代的到来。

1 科研智能的概念及内涵

科研智能是科学智能概念的延伸和扩

展，主要包括两方面内容：一是聚焦人工智能在基础科学研究领域的应用，加速知识发现，解析内在规律和结构，如数学猜想证明、物理规律发现等；二是强调人工智能在产业研发领域的应用，优化技术研发流程，发明新方法、新工具和新产品，如新材料研发、新药研制等。科研智能代表了一种新的科技创新范式，有望全面加速基础研究和产业研发的进程，并缩短两者之间的转化周期，推动科技成果在工业界的产业化和规模应用。本文采用“科研智能”这一概念，旨在强调人工智能对科学研究和产业研发两个领域的赋能作用^[6]。

科研智能的核心特征包括 3 个方面。一是数据驱动。人工智能擅长处理海量、高维度数据，并提取其中的模式和关联，为各领域科研人员提供统一的工具和方法论。例如，AI 可通过分析材料的组成-结构-性能关系，从数据库中筛选潜在的新型高性能材料。二是模型智能。人工智能可以高效模拟和优化复杂系统，尤其适用于多变量、多尺度或非线性特征的系统。例如，AI 可以模拟气候变化，预测长中短期天气变化。三是自动化与优化。人工智能可以实现实验设计、数据分析和实验验证的自动化。例如，强化学习可用于设计最优实验方案，减少试验次数并提高效率；AI 与机器人等技术相结合可以设计自动化、智能化、高通量且能实现“干湿闭环”的实验验证平台，大幅提高实验效率。

科研智能被视为继实验范式、理论范式、计算范式和数据驱动范式之后的“第五种科研范式”，它与传统科研范式在研究目标和流程上有继承关系，同时极大拓展了科研的能力边界，使科研工作朝着高效化、自动化和跨学科方向发展。与传统科研相似的是，科研智能依旧关注对自然规律的探索、科学问题的解决和技术创新，

沿用并完善了“提出假设—设计实验—采集数据—分析验证”的基本流程，它的核心技术发展也依赖统计学、计算机科学等传统学科的理论贡献。与传统科研相比，科研智能更善于处理规模庞大、类型多样且高度异构的复杂数据，挖掘隐藏的规律；在深度学习与强化学习等方法的支持下，AI 可对复杂系统进行有效模拟，并进行高效预测；通过全流程自动化和优化，科研效率与精度大幅提升。人工智能既是工具，也是科研合作者，未来的人机协同模式将显著增强跨学科资源的整合与共享，进一步激发跨领域协同创新。

科研智能与普通商业 AI 也有普遍的联系和明显的区别。一方面，两者拥有技术共通性，互相促进。深度学习、强化学习等算法在科研与商业活动中得到广泛应用。例如，生成式 AI 技术（如 Transformer、扩散模型）的出现带动了科研领域 AlphaFold 模型的成功，以及商业领域 ChatGPT 等产品的火爆。科研智能提供的新算法与模型理论能够反哺商业应用，商业应用场景又可推动科研智能在工程化层面的发展。另一方面，两者又有许多不同。科研智能更专注于科学与技术难题的攻关，数据复杂、技术前沿，成果转化周期长；普通商业 AI 致力于提升商业效率，数据规模大但应用直接，追求短期效益。其中，普通商业 AI 使用的数据主要是人类活动产生的数据，尤其是结构化数据和常规非结构化数据（如文本、图片和视频等），数据量大但结构化程度较高；相比之下，科研智能处理的数据往往源自更广泛的自然现象，如生物医药中的 DNA、RNA 和蛋白质序列，化学材料中的分子结构和晶体结构等，这些数据通常具有高维度、跨模态、多尺度和强约束性。在此背景下，科研智能的发展有望大幅扩展人类对科学和技术的认知边界。

2 科研智能政策支持

2.1 科研智能政策的分类框架

随着人工智能深入赋能科研活动，各国政府相继出台政策来促进科研智能发展。本文提出一种新的政策分类框架，按政策取向划分为3类，分别为立法导向型、资源导向型和场景导向型，涵盖了欧美与我国的主要政策实践及其特征。

立法导向型政策强调通过法律法规和伦理准则来规范和引导科研智能发展，建立明确的制度框架和价值导向，以确保科研活动的安全、可信和负责任。欧盟是典型代表，2024年8月生效的《人工智能法案》作为全球首部综合性AI法律，对科研人工智能系统给予合规豁免，以平衡监管与创新^[7]。同年发布的文件《科研领域负责任使用生成式人工智能指南》进一步细化了科研伦理指引^[8]。同时，欧盟也开展了大量预研工作支持科研智能政策出台。美国尚未出台AI统一立法，但通过《人工智能权利法案蓝图》等软法框架提出伦理准则，并通过现有法律体系规范科研智能。2023年10月，美国发布《关于安全、可靠、可信地开发和人工智能的行政命令》，要求体系化推动科研智能工作^[9]。2024年4月，美国参议院提出立法《人工智能创新未来法案》，强调构建AI安全框架，并支持科研应用。我国则通过一系列法规文件构建多层次的治理体系：2021年修订的《中华人民共和国科学技术进步法》为AI赋能科研提供法律依据；2023年发布的《生成式人工智能服务管理暂行办法》对科研用途AI作出豁免；最新版《科研诚信规范手册》也对生成式AI作出合规指引。此类政策体现了制度化、责任化的发展思路^[10]。

资源导向型政策侧重于通过投入各类资源来支撑科研智能的发展，包括资金支持、基础设施建设和人才培养等。美国于2024年年初正式启动国家人工智能研究资源（NAIRR）计划试点，目标是建立全国性共享AI科研基础设施，推动AI研发“民主化”，截至2025年6月已支持400余个项目，其中25%的项目聚焦科研智能^[11]。2024年年中，美国能源部提出面向科学、安全和技术的人工智能前沿倡议（FASST Initiative），依托其下属17个国家实验室的资源（大科学装置、海量科研数据、AI计算设施、4万余名科研人员等），构建面向科研、能源和安全的科学AI系统^[12]。欧盟通过地平线欧洲（Horizon Europe）、欧洲研究委员会、欧洲创新委员会和欧洲伙伴关系（European Partnerships）提供多层次支持，涵盖基础研究、前沿探索、技术创新及应用落地。我国则在“科技创新2030-重大项目”“国家重点研发计划”等框架下持续投入。2023年启动的“人工智能驱动的科学驱动”专项重点支持基础科学与药物、新材料等关键领域的创新。此类政策不仅关注成果数量，更着眼于长期基础投入和高端人才培养，厚植发展根基。

场景导向型政策围绕特定应用场景或战略任务设定研究目标，强调科研与现实需求相结合。美国能源部2021年启动AI用于地球系统可预测性计划，利用AI增强对地球系统（气候、大气、水文等）的预测能力^[13]。美国国立卫生研究院2022年启动桥接AI计划，加速AI在生物医学和行为科学研究中的广泛应用^[14]。我国多部门于2022年发布《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》，提出打造高水平科研场景，突破基础科学及新药创制、新材料研发等重点领域技术瓶颈。欧盟也通过地平线计划、

数字欧洲计划等设定多领域场景任务，推动 AI 赋能医疗、农业、气候等领域。此类政策促进科研与场景需求联动，伴随跨部门协作，有助于成果转化。

上述 3 类政策并非孤立存在的，美国侧重“资源+场景”导向，以宽松监管鼓励主动创新；欧盟体现“立法+资源”，强调伦理规范与合作自助并重；我国则为“资源+场景+战略引导”，强化顶层设计并引导科研资源优化配置。该分类框架有助于理解各国政策逻辑，为政策成效评估提供参考。

2.2 不同科研智能政策范式的优劣分析

基于上述分类框架，本文从科研效率和科研质量两个维度，对比 3 类政策的优势与不足。考虑到政策效果评估是一项长期且复杂的工作，本文主要进行定性分析，详见表 1。

2.3 技术演进对科研政策的新需求

随着大模型和自动化科研系统等兴起，科研政策也需要适应技术演进。首先，

政策应支持科研数据与计算基础设施建设，包括搭建科研数据平台、高性能计算中心，以及完善科研数据流通机制。其次，需完善科研智能伦理与治理框架。大模型与自动化科研系统可能引发偏见、透明性与诚信问题，政策需建立伦理规范与技术审查机制。最后，强化跨机构与国际协作机制势在必行。国内方面，应促进高校、科研机构、企业间的联合攻关；国际方面，在气候、健康等全球议题上应推动共享数据与联合研究。综上，科研智能政策的持续优化需立足基础、回应前沿、兼顾规范与灵活，才能在保障伦理与安全的前提下释放科研创新潜能。

3 科研智能关键技术

人工智能正深入赋能科学研究和产业研发的各个场景，并持续融入科研的全流程，加快创新速度。本文综述科研智能领域的关键技术，立足基础通用视角，将其分为科研模型、科研智能体和自动化实验室，分别阐述各类技术的核心内容与发展趋势。

表1 3类科研智能政策的优劣分析

政策类型	内容重点	对科研效率的影响	对科研质量的影响
立法导向	法规与伦理	优点:提供明确规则,减少低质或违规研究的资源浪费; 缺点:手续烦琐增加合规成本,科研速度放缓	优点:确保研究遵循高标准和社会价值,提升成果可靠性和公信力; 缺点:可能抑制创新尝试,降低颠覆性成果出现的概率
资源导向	资源投入与平台建设	优点:可迅速扩大研究规模,增加论文产出和项目数量; 缺点:边际效益递减,若分配低效,可能造成资源浪费	优点:可支持精尖研究和大项目攻关,潜在催生重大突破; 缺点:资源富集环境下易滋生平庸研究、重复研究
场景导向	应用需求牵引	优点:加快成果落地转化,缩短从实验室到应用的周期; 缺点:资源过度集中于限定场景,其他领域进展放缓	优点:以实际需求检验成果,有助于产出实用创新; 缺点:可能忽视基础原理探索,研究的通用性及长远性受限

3.1 科研模型

科研模型是用于科研领域的人工智能模型和算法集合，主要面向科研问题的数据分析和知识发现，可分为科研大语言模型、科研专用模型和科研多模态模型。

科研大语言模型经过科研数据强化训练，先通过自监督预训练学习科研知识，部分科研 LLM 经指令微调与人类反馈强化，具备科研场景的阅读理解、知识问答、推理和代码生成能力。通用科研 LLM（如 Galactica、SciGLM）可用于综述研究和跨领域讨论^[15-16]；专用科研 LLM（如生命科学领域的 BioGPT、Med-PaLM 2）可用于结果解读和病例分析^[17-18]，化学材料领域的 ChemDFM、ChemLLM 则可辅助实验设计及材料筛选^[19-20]。

科研专用模型针对特定科研问题，使用专业数据训练，专注高精度模拟物理过程或解决专业任务，旨在代替或者强化传统理论模型。此类模型通常融合物理先验、大规模数据驱动以及 Transformer/图网络/神经算子，以端到端学习“直接映射”代替传统科学计算“求解方程”。代表性进展有 AlphaFold 模拟蛋白质折叠精度堪比实验测定、盘古气象模型首次超越欧洲中期天气预报中心数值模式的预测精度^[21]。

科研多模态模型能够同时处理和融合多种数据模态，统一表示学习框架实现联合分析和决策，属于前沿方向。代表性进展包括：生物医药领域，BiomedGPT 和 Biomedical CLIP 使用医疗影像和文本跨模态训练^[22-23]；化学材料领域，MultiMat 融合晶体结构、电子密度谱、文本信息等模态，并支持模态扩展^[24]；地球科学领域，SkySense 融合多源遥感数据，泛化多种环境任务^[25]。

未来，科研专用模型有望在更多领域取得类似 AlphaFold 的突破，多模态大模型将进一步融入分子结构等新模态，科研

通用基础大模型也将加速发展。此外，可解释性将成为研究重点。

3.2 科研智能体

科研智能体是能自主执行科研活动的人工智能系统，结合语言理解、规划决策、工具使用和环境交互等能力，模拟或辅助人类科学家的研究过程。智能体通常由多个组件和工具组成，具备记忆、推理和行动能力，能够执行从辅助科研任务到自主循环实验等多种任务。

科研智能体的构建除科研大模型外，还需如下技术支撑：自动推理与工具使用，通过思维链输出决策序列，通过接口调用外部工具；自动实验设计与执行，调用自动化实验室接口执行实验步骤；多模态交互，处理文本、图像、谱图、分子结构等数据；多智能体协同，利用多个不同专长的智能体合作解决复杂课题。

近年来涌现出诸多科研智能体系统，既有通用型，也有专用型。例如，2024 年 Sakana AI 提出的 AI 科学家系统（The AI Scientist），首次实现了自主开放式科研循环^[26]；2025 年谷歌开发的 AI 联合科学家系统（AI Co-Scientist）可协助假设生成与实验设计，并在生物医药领域验证了其成效^[27]；化学材料领域的 ChemCrow、PolySea 可实现材料设计等任务^[28-29]。

当前，基于 LLM 的科研助理、多智能体协作和人机协同都是研究热点。未来，科研智能体与自动化实验室将深度融合，形成“设计-执行”一体化科研体系，加速自动化科研发现。

3.3 自动化实验室

自动化实验室高度依赖机器人、自动控制和信息化系统执行实验操作，其功能

覆盖实验设备操作、样品处理、数据测量与记录，以及与智能系统交互。

自动化实验室依赖多学科技术融合，包括自动化硬件系统、实验调度管理系统与数据管理平台。硬件系统含机械臂、移动机器人以及自动化实验装置，实现高精度、高通量操作；调度管理系统通过流程描述语言、任务规划算法优化实验步骤；数据管理平台支持标准化数据格式、自动记录、开放共享与远程协作。

自 2020 年以来，全球出现一批标志性自动化实验室系统：利物浦大学研制出首个移动机器人化学家，其可以自主开展各种实验^[30]；中国科学技术大学机器化学家平台小来可覆盖文献阅读、实验设计到合成测试的全流程^[31]；劳伦斯伯克利国家实验室 A-Lab 平台实现材料合成与测试自动闭环^[32]；2024 年多伦多大学联合全球多个团队首次实现跨时区、跨设施、云端协同的自动化实验室协作网络^[33]。

当前该领域的发展聚焦硬件层面的全流程自动化整合。未来还需要提升系统泛化能力，扩展适用领域；与科研模型和智能体深度融合提升智能水平；构建不同学科、不同地点之间的泛在实验室网络，提升解决复杂科研问题的效率。

4 科研智能典型应用场景

近年来，人工智能技术与科研的深度融合正在重塑科学研究与产业创新的范式。在基础科学领域，AI 通过多尺度建模、高通量计算与数据驱动分析等技术，突破了传统研究方法的效率瓶颈，催生出一系列颠覆性成果。从蛋白质结构预测到新型材料设计，从细胞行为解析到药物分子生成，人工智能不仅加速了科学发现的进程，更

通过“算法创新-物理约束-数据增强”的协同创新开辟了全新的研究路径。与此同时，在产业研发领域，AI 技术与自动化实验平台、机器人系统的深度融合，正在构建“智能设计-自主验证-快速迭代”的研发闭环，显著提升医药、材料等战略产业的创新效率。本节旨在梳理人工智能在基础科学与产业研发典型领域的进展，分析其典型应用场景及产业动向，并探讨这些进展对未来科研与产业发展的深远意义。

4.1 基础科学领域

生物领域，人工智能在生命科学中的应用已经深入多个层面（从分子、细胞、系统到个体），为研究人员提供了强有力的分析和预测工具^[34]。分子层面，近年来最重要的突破是 DeepMind 开发的 AlphaFold 系列模型。2020 年 AlphaFold2 首次实现了接近实验精度的蛋白质结构预测，可直接从氨基酸序列推测蛋白质的三维结构，这是该领域取得的巨大进步。随后两年，DeepMind 预测并公开了约 100 万个物种的 2.14 亿种蛋白质结构，几乎涵盖了地球上已知的所有蛋白质。2024 年，AlphaFold3 进一步扩展至所有生命分子结构及相互作用，包括蛋白质、DNA、RNA 与配体等，预测精度和适用范围显著提升^[35]。AlphaFold 相关工作于 2024 年获得诺贝尔化学奖，对生命科学研究产生了深远影响。细胞层面，人工智能可以模拟细胞内的生物过程（如信号传导、代谢途径等），预测细胞行为和响应模式，帮助人们理解疾病的发生机制及治疗方法。系统层面，人工智能可以处理和分析大规模且复杂的生物数据集（如基因组学、转录组学等），更深入地理解生物系统的运作机制。机体层面，通过对大规模个体数据（基因

组、临床数据等)的整合,人工智能为个性化医学研究和健康管理开辟了新途径。

化学领域,人工智能在分子设计、反应预测、材料设计和药物设计等方面的应用,正在深刻影响着化学研究与产业实践^[36]。其中,生成式模型和大语言模型推动了新分子的智能生成与优化,还可实现反应条件与结果的高准确率预测,结合自动化平台推动反应规划,分子设计和反应预测领域的进展加速了材料设计和药物设计领域的突破。以材料设计为例,2023年11月,DeepMind推出基于图神经网络的工具GNoME,其成功预测出220万种新型晶体结构,其中38万种被证实具有稳定性,这些发现使已知稳定无机晶体数量增加了8倍,材料稳定性预测准确率也从传统方法的不足50%提升至80%^[37]。此外,DeepMind与劳伦斯伯克利国家实验室合作,利用自主机器人实验室成功合成41种预测材料(成功率70.7%),全球各地研究团队独立合成的数量达736种,有望大幅加速新能源材料、半导体材料等的研发进展。2025年1月,微软发布生成式AI材料设计工具MatterGen,推动了材料研发从“筛选”到“创造”的转变,其可根据应用需求直接生成所需化学、机械、电子或磁性属性,以及满足不同约束条件的新材料^[38]。MatterGen生成材料的新颖性和稳定性比传统方法高两倍以上,且更接近能量最小值。

4.2 产业发展领域

医药研发领域,人工智能驱动的药物研发正在快速崛起,推动制药行业的快速发展。产业应用层面,AI的主要应用是药物靶点发现、药物分子设计与筛选、药物分子合成等^[39]。据波士顿咨询公司2024年

4月的研究^[40],全球排名前20的药企均已开展AI+医药研发,其中多采用与AI生物医药创业公司合作的模式。数据显示,截至2023年年底,有67种AI驱动的药物分子进入临床试验,过去10年的年复合增长率超60%。其中,24种分子完成临床一期实验,成功率为87.5%,显著高于行业平均的40%~65%;10种分子完成临床二期实验,成功率达40%,与行业平均水平相当,展示出巨大的应用潜力。市场层面,全球AI医药市场规模快速增长,我国起步较晚,但增长迅速且未来潜力巨大。根据MedMarket Insights和智药局等机构的数据^[41-42],2023年全球AI医药市场规模达12.93亿美元,我国市场规模约10.24亿元;预计到2031年,分别增长至85.02亿美元和116.31亿元,年复合增长率分别为27.3%和35.2%。在投融资方面,中美领衔全球投融资,美国优势显著,且头部公司吸引了大部分资本,我国市场融资集中在早期阶段。据国家药品监督管理局研究,2024年,美国融资事件64起,融资金额全球占比约80%,我国融资事件37起,融资金额全球占比约10%;排名前10的公司(均为美国公司)融资总金额达26.9亿美元,占全年总额的46%。

材料研发领域,人工智能正在深刻改变材料研发的传统模式,其应用场景涵盖材料研发的多个环节。产业应用层面,AI主要应用场景包括新材料的发现与设计、实验自动化与高通量验证等。多家企业推出商业化AI材料研发平台。如2023年3月,晶泰科技推出UpChemist.AI平台,结合理论物理、量子化学、人工智能和机器人技术,聚焦新材料的发现与优化;与广东省科学院化工研究所合作仅4个月,就成功设计出两款生物基表面活性剂,可代替市面主流的石油基产品,并获得权威实验

验证^[43]。2024 年 11 月，英伟达发布化学材料研发平台 NVIDIA ALCHEMI，通过生成式 AI 模型和 AI 代理模型，该平台可将材料的长期模拟稳定性评估速度提高 100 倍，促进锂电池等新能源材料的快速迭代^[44]。市场层面，据 Dimension Market Research 和头豹研究院等统计^[45-46]，2024 年，全球 AI 材料市场规模约为 12 亿美元，我国市场规模约为 14.7 亿元；到 2028 年，两者将分别增长至 35.23 亿美元与 34.17 亿元，年复合增长率分别为 30.9% 和 23.48%。全球来看，专注于 AI 驱动材料研发的创企数量仍然较少，投融资案例集中于少数头部项目，单笔金额通常较高，整体处于初创探索阶段。在我国，随着国家政策扶持与资本活跃度提升，AI+ 材料研发领域正迎来高速发展，形成“北上广”集聚的产业格局。

5 科研智能发展趋势

展望未来，科研智能主要呈现如下发展趋势。

一是科研范式的系统性变革。人工智能正逐步嵌入科研的全流程，从假设提出、实验设计、数据采集、分析到结果验证，AI 正在从辅助工具变为核心驱动力。

二是技术创新推动高速发展。物理引导的机器学习将提升模型的可解释性，通过将领域知识（如物理定律、守恒方程等）与数据驱动模型相结合，进一步增强模型的可解释性、泛化能力和鲁棒性。多模态 AI 技术有望成为科研的“全局分析工具”，通过综合分析不同来源的数据（实验、计算、文献、数据库等），帮助科研人员挖掘深层次的规律和洞察。同时，融合智能化和机器人技术的自主实验平台将在科研中

扮演关键角色，通过有效整合计算模拟与实验验证构成的“干湿闭环”，双轮驱动的科研模式将显著提升科研效率、拓展科研空间。

三是跨学科融合重塑科研生态。人工智能作为通用技术，能更紧密地连接数学、物理、生物、化学、计算机等多个学科，推动跨学科团队的协同研究，催生一批新的交叉学科领域。

四是应用场景持续拓展。科研智能在生物、化学等基础科学领域取得突破性成果后，还将更多地应用于医药、材料、能源等产业研发场景，形成理论创新与产业实践的高效衔接，进一步加速科研成果的落地转化。

6 结束语

本文系统梳理了科研智能的发展历程、政策支持及其在基础科学和产业研发领域的典型应用，揭示了人工智能作为“第五种科研范式”对科学研究和产业创新的深远影响。从 AlphaFold 在蛋白质结构预测中的突破，到生成式 AI 在材料设计和药物研发中的广泛应用，科研智能不仅显著提升了科研效率，还开辟了全新的研究路径和技术范式。政策层面，中国、美国和欧盟通过立法、资金支持和项目引导，构建了各具特色的科研智能发展框架，为全球科研智能的快速发展提供了有力支撑。未来，随着物理引导的机器学习、多模态 AI 技术和自主实验平台的进一步发展，科研智能将更加深入地嵌入科研全流程，推动跨学科融合和科研成果的快速转化。然而，科研智能的发展也面临数据质量、模型可解释性、伦理风险等挑战，需要学术界、产业界和政策制定者共同努力，构建更加

开放、协作和负责任的科研生态。展望未来，科研智能有望在更多领域取得突破性进展，为人类探索自然规律和解决全球性挑战提供更强大的工具。

参考文献：

[1] Tansley S, Tolle K M. The fourth paradigm: data-intensive scientific discovery[M]. Redmond: Microsoft Research, 2009.

[2] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold[J]. Nature, 2021, 596(7873): 583-589.

[3] U.S. Department of Energy. AI for science[R]. 2020.

[4] 程学旗, 梅宏, 赵伟, 等. 数据科学与计算智能: 内涵, 范式与机遇[J]. 中国科学院院刊, 2020, 35(12): 12.

Cheng X Q, Mei H, Zhao W, et al. Data science and computing intelligence: concept, paradigm, and opportunities[J]. Bulletin of the Chinese Academy of Sciences, 2020, 35(12): 12.

[5] 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. AI4R: the fifth scientific research paradigm[J]. Bulletin of Chinese Academy of Sciences, 2024, 39(1): 1-9.

[6] 中国人工智能产业发展联盟. 科研智能(AI4R&D)——人工智能驱动的研发新范式报告[R]. 2024.

Artificial Intelligence Industry Alliance. AI for R&D (AI4R&D): a new paradigm of AI-driven research and development[R]. 2024.

[7] European Union. Regulation (EU) 2024/1689 on Artificial Intelligence (AI Act)[S]. Brussels: Official Journal of the EU, 2024.

[8] European Research Area Forum. Guide-

lines on the responsible use of generative AI in research[Z]. 2024.

[9] Biden J R. Safe, secure, and trustworthy development and use of artificial intelligence[Z]. 2023.

[10] 国家自然科学基金委员会. 科研诚信规范手册[Z]. 2023.

National Natural Science Foundation of China. Handbook of research integrity guidelines[Z]. 2023.

[11] National Science Foundation. National artificial intelligence research resource (NAIRR) Pilot[Z]. 2024.

[12] U.S. Department of Energy. Frontiers in artificial intelligence for science, security and technology (FASST)[Z]. 2024.

[13] Sagoff J. Predicting the future of the Earth with artificial intelligence[Z]. (Lemont): Argonne National Laboratory, 2021.

[14] National Institutes of Health. Bridge2AI program launch notice[Z]. 2022.

[15] Taylor R, Kardas M, Cucurull G, et al. Galactica: a large language model for science[EB]. arXiv preprint, 2022, arXiv: 2211.09085.

[16] Zhang D, Hu Z N, Zhou B S N, et al. SciGLM: training scientific language models with self-reflective instruction annotation and tuning[EB]. arXiv preprint, 2024, arXiv: 2401.07950.

[17] Luo R Q, Sun L A, Xia Y C, et al. BioGPT: generative pre-trained transformer for biomedical text generation and mining[J]. Briefings in Bioinformatics, 2022, 23(6): bbac409.

[18] Singhal K, Tu T, Gottweis J, et al. Toward expert-level medical question answering with large language models[J]. Nature Medicine, 2025, 31(3): 943-950.

[19] Zhao Z H, Ma D, Chen L, et al. ChemDFM: a large language foundation model for chemistry[EB]. arXiv preprint, 2024, arXiv: 2401.14818.

- [20] Zhang D, Liu W, Tan Q, et al. ChemLLM: a chemical large language model[EB]. arXiv preprint, 2024, arXiv: 2402.06852.
- [21] Bi K F, Xie L X, Zhang H H, et al. Accurate medium-range global weather forecasting with 3D neural networks[J]. Nature, 2023, 619(7970): 533–538.
- [22] Zhang K, Zhou R, Adhikarla E, et al. A generalist vision-language foundation model for diverse biomedical tasks[J]. Nature Medicine, 2024, 30(11): 3129–3141.
- [23] Zhang S, Xu Y B, Usuyama N, et al. BiomedCLIP: a multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs[EB]. arXiv preprint, 2023, arXiv: 2303.00915.
- [24] Moro V, Loh C, Dangovski R, et al. Multimodal foundation models for material property prediction and discovery[J]. Newton, 2025, 1(1): 100016.
- [25] Guo X, Lao J W, Dang B, et al. SkySense: a multi-modal remote sensing foundation model towards universal interpretation for earth observation imagery[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 27662–27673.
- [26] Yamada Y, Lange R T, Lu C, et al. The AI scientist-v2: workshop-level automated scientific discovery via agentic tree search[EB]. arXiv preprint, 2025, arXiv: 2408.06292.
- [27] Gottweis J, Weng W H, Daryin A, et al. Towards an AI co-scientist[EB]. arXiv preprint, 2025, arXiv: 2502.18864.
- [28] Bran A M, Cox S, Schilter O, et al. Augmenting large language models with chemistry tools[J]. Nature Machine Intelligence, 2024, 6(5): 525–535.
- [29] Qiu H, Zhao J, Jing E, et al. Introducing PolySea: an LLM-based polymer smart evolution agent[EB]. 2025.
- [30] Burger B, Maffettone P M, Gusev V V, et al. A mobile robotic chemist[J]. Nature, 2020, 583(7815): 237–241.
- [31] Zhu Q, Zhang F, Huang Y, et al. An all-round AI-chemist with a scientific mind[J]. National Science Review, 2022, 9(10): nwac190.
- [32] Szymanski N J, Rendy B, Fei Y X, et al. An autonomous laboratory for the accelerated synthesis of novel materials[J]. Nature, 2023, 624(7990): 86–91.
- [33] Strieth-Kalthoff F, Hao H, Rathore V, et al. Delocalized, asynchronous, closed-loop discovery of organic laser emitters[J]. Science, 2024, 384(6697): eadk9227.
- [34] 李丹丹, 马征远, 张博文, 等. 人工智能驱动生命科学研究的态势[J]. 生命科学, 2024, 36(1): 122–128.
- Li D D, Ma Z Y, Zhang B W, et al. The development trend of artificial intelligence driven life science research[J]. Chinese Bulletin of Life Sciences, 2024, 36(1): 122–128.
- [35] Abramson J, Adler J, Dunger J, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3[J]. Nature, 2024, 630(8016): 493–500.
- [36] Yang L, Guo Q L, Zhang L J. AI-assisted chemistry research: a comprehensive analysis of evolutionary paths and hotspots through knowledge graphs[J]. Chemical Communications, 2024, 60(55): 6977–6987.
- [37] Merchant A, Batzner S, Schoenholz S S, et al. Scaling deep learning for materials discovery[J]. Nature, 2023, 624(7990): 80–85.
- [38] Zeni C, Pinsler R, ZüGner D, et al. A generative model for inorganic materials design[J]. Nature, 2025, 639(8055): 624–632.
- [39] 万宇, 郑维恒, 蒋阅, 等. 人工智能在制药行

业的应用进展[J]. 药学进展, 2023, 47(10): 741-750.

Wan Y, Zheng W H, Jiang Y, et al. Ap - plication progress of artificial intelli - gence in the pharmaceutical industry[J]. Progress in Pharmaceutical Sciences, 2023, 47(10): 741-750.

[40] Kp Jayatunga M, Ayers M, Bruens L, et al. How successful are AI-discovered drugs in clinical trials? A first analysis and emerging lessons[J]. Drug Discovery Today, 2024, 29(6): 104009.

[41] 智药局. 2023年AI制药行业报告[R]. 北京: 智药局, 2023.

Zhiyaoju. 2023 AI-driven pharmaceutical industry report[R]. Beijing: Zhiyaoju, 2023.

[42] 智药局. 2024全球AI制药融资强势回春[EB]. 北京: 智药局, 2025.

Zhiyaoju. 2024 Global AI-driven pharma - ceutical financing strong rebound[EB]. Beijing: Zhiyaoju, 2024.

[43] 张馨怡. AI+机器人破局,河套这家企业4个月研发两款生物基表面活性剂[EB]. 2023.

Zhang X Y. AI+robotics breakthrough: this Hetao-based company developed two bio-based surfactants in four months[EB]. 2023.

[44] Nvidia Corporation. From algorithms to atoms: NVIDIA ALCHEMI NIM cata - lyzes sustainable materials research for EV batteries, solar panels and more[EB]. 2024.

[45] Dimension Market Research. Generative AI in material science market: global industry outlook, key companies, trends and forecast 2024-2033[R]. 2024.

[46] 头豹研究院. AI+材料:AI重塑材料科学,驱动性能优化与成本革命[Z]. 2024.

LeadLeo Research Institute. AI+ mate - rials: AI reshaping materials science, driving performance optimization and cost revolution[Z]. 2024.

作者简介



董昊（1987-），男，中国信息通信研究院工程师，人工智能研究所平台与工程化部副主任，主要研究方向为科研智能、人工智能基础设施、人工智能工程化等。



魏凯（1981-），男，中国信息通信研究院正高级工程师，人工智能研究所所长，主要研究方向为人工智能等。

收稿日期: 2025-06-24

通信作者: 董昊,donghao@caict.ac.cn