人工智能将推动材料基因组技术加速发展

黄 河，陈宏生
（中国科学技术部，北京 100862）

摘 要：本文通过对美国材料基因组计划出台、实施情况进行梳理分析，发现人工智能在美国材料基因组技术中的应用正在成为研究热点，其中机器学习、深度学习等算法是关键。人工智能在解决材料基因组计划发展瓶颈问题方面非常有效，有可能会推动材料基因组技术研究加速发展，但也面临不少挑战。在此基础上，提出加快推动我国材料基因组技术研究的有关建议。

关键词：美国；材料基因组计划；人工智能；机器学习；深度学习
中图分类号：F062.9 **文献标识码**：A DOI：10.3772/j.issn.1009-8623.2019.11-12.008

材料是所有产业的基础和先导，材料从研发到投入市场的时间跨度很长，其关键在于长期以来材料研发过度依赖科学直觉与试错式的实验经验积累，且制备过程漫长和充满变数。传统的材料研发模式主要是以实验为主的“试错法”，效率低，从新材料的最初发现到最终工业化应用一般需要10~20 年。如目前移动电子设备所用的锂电池，从20 世纪 70 年代中期的实验室原型到 90 年代实现应用，前后花了近 20 年时间[1]，但是至今还不能完全、充分地应用到电动汽车上。变革材料的研究与开发方式、提高材料从发现到应用的速度成为世界各国共同的追求[2]。为了加快材料的研发，2011年6 月美国设立“材料基因组计划（Material Genome Initiative，MGI）”，借鉴生物基因组的理念，利用先进的计算手段和数据方法，以集成化的“多尺度计算 – 高通量实验 – 数据库技术”为核心，旨在达到新材料研发周期缩短一半、研发成本降低一半的目的，最终支撑先进制造和高新技术的发展[1]。自从美国率先启动材料基因组计划以来，中国、欧洲和日本等也先后开展了类似的研究计划，争取在新一轮的材料革命性发展中抢占先机[3]。

近来，美国标准技术研究所、能源部先后组织多场研讨会，材料基因组计划执行秘书、美国标准技术研究所材料测量实验室材料基因组计划主任詹姆斯 · A. 沃伦（James A. Warren，作为美国国家科技委员会委员参与制定《具有全球竞争力的材料基因组计划》白皮书）等发表文章和演讲，围绕人工智能（AI）在材料基因组技术中的应用进行讨论，提出了很多前瞻性观点，值得我国关注和借鉴。

1 美国材料基因组计划实施概况

2011 年 6 月 24 日，美国时任总统奥巴马在卡耐基 · 梅隆大学发表《先进制造业伙伴关系》（Advanced Manufacturing Partnership，AMP）主题演讲，宣布启动美国“材料基因组计划”，设定目标为“将先进材料的发现、开发、制造和使用的速度提高1 倍”。该计划由美国国家科技委员会（NSTC）下设的材料基因组计划小组委员会（SMGI）具体管理，美国能源部、国防部、国家科学基金会、标准技术研究所、食品和药物管理局、国家科技委员会、白宫科技政策办公室、白宫总统预算管理办公室、国家航空航天管理局、陆军、空军、国防部先进技术研究所、网络与信息技术研发计划（NITRD）、国家纳米技术计划、国家核安全管理局等 18 个政

第一作者简介：黄河（1976—），男，一级调研员，主要研究方向为科技管理，科技政策研究。
收稿日期：2019−10−23

府、军队和科研机构参与该计划。

2011 年 6 月美国国家科技委员会和白宫科技政策办公室发布《具有全球竞争力的材料基因组计划》白皮书，阐述了材料创新基础设施的 3 个平台，即计算工具平台、实验工具平台和数字化数据（数据库及信息学）平台，如图 1 所示，明确指出材料基因组计划不仅仅要开发快速可靠的计算方法和相应的计算程序，而且要开发高通量的实验方法对理论进行快速验证，为数据库提供必需的输入，建立普适可靠的数据库和材料信息学工具，以加速新材料的设计和使用 [4]。白皮书明确了材料基因组计划的 3 个目标：一是建立新的材料创新基础设施；二是利用先进材料实现国家目标；三是培养新一代材料人才。这 3 个目标将使得美国研发、制造和应用先进材料的速度比现在至少快一倍、成本只有现在的一小部分，支撑美国制造业的全球竞争力 [4]。

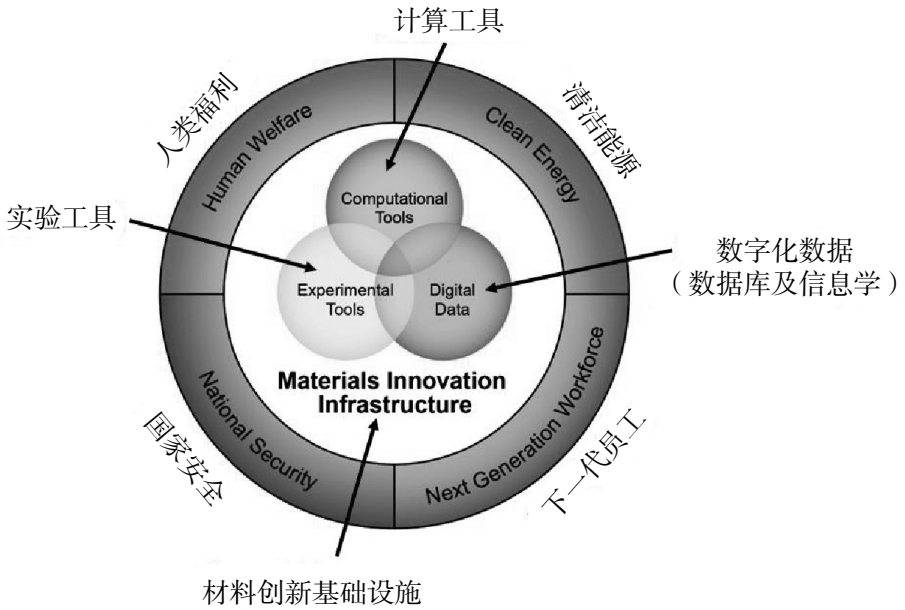

图 1　材料基因组计划材料创新基础设施 [5]

2014 年 12 月美国国家科技委员会颁布《材料基因组计划战略规划》，其核心内容仍然是建立高通量材料计算方法、高通量材料实验方法和材料数据库，但其中一个重要的改变是确立了“新材料的发现从设计开始”的原则，即注重从原子与分子层面上认识、设计和计算新材料，并通过数据库收集已有材料的结构与性能的相关性，从而指导新材料的设计和开发。该规划进一步明确了材料基因组计划的目的：一是通过高通量筛选新材料，加快材料的研发进程；二是转变新材料研发范式，节省人力、物力；三是加快人类对材料本质与规律的认识，再通过所认识的材料本质与规律指导新材料的设计、制备与检测等，从而加深对材料本质与规律的认识；四是建造可靠的材料基因组数据库（新材料大数据），实现资源共享，加速新材料的发现和应用（见图 2）[3, 6]。

2012 年 5 月 14 日，美国总统执行办公室发布材料基因组计划进展简报，能源部、国防部、国家科学基金会、标准技术研究所 4 个联邦机构共设立了 9 个研发项目，哈佛大等 34 所大学、洛克希德・马丁和通用电气等 33 家公司、阿贡国家实验室、伯克利国家实验室等高校和机构陆续加入该计划 [7]。2013 年 6 月 24 日，美国总统执行办公室发布材料基因组计划进展简报，麻省理工学院、佐治亚理工大学、威斯康星麦迪逊大学、材料研究学会等高校和机构宣布加入该计划，能源部、国防部、国家科学基金会、标准技术研究所等联邦机构宣布投入更多资金用于该计划的研发 [8]。

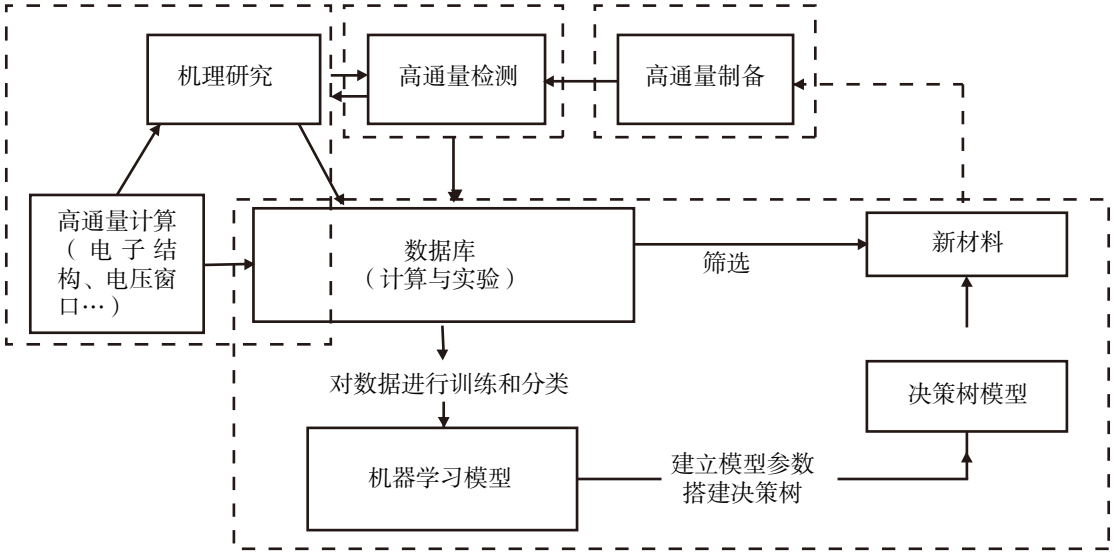

图 2　材料基因组计划总体研发路线图 [3]

2014 年 6 月 19 日，白宫科技政策办公室发布材料基因组计划进展简报，美国政府过去 3 年在该计划创新基础设施方面共投入 2.5 亿美元，并宣布政府将继续扩大投资，不少私营部门和学术合作伙伴承诺加入和支持该计划 [9]。2016 年 8 月 2 日，白宫科技政策办公室发布《材料基因组计划五年成就和技术亮点》报告，指出过去 5 年来，联邦政府共投入 5 亿美元支持该计划。该计划在引领材料研究方式转变，整合实验、计算和理论，为材料界配备先进的工具和技术，使数字化数据可访问，培养世界一流的材料科学人才队伍等方面取得进展，材料基因组计划已经引发了新材料研究、开发和应用范式的转变 [10, 11]。

2013 年 9 月、2015 年 1 月、2016 年 1 月，美国自然基金会、能源部、标准技术研究所分别牵头召开了 3 届材料基因组计划首席调查员年会，集中对该计划研究项目进行深入讨论和交流 [12, 13]。2019 年 3 月美国自然基金会、能源部、标准技术研究所牵头召开了第 4 届材料基因组计划首席调查员会议，讨论如何通过实施战略来实现该计划制定的目标，包括减少将新材料推向市场所需的成本和时间，从而为更广泛的材料研究提供指导。研讨会的主要目的之一就是要加强材料基因组计划的数据驱动研究，推动应用人工智能和机器学习技术来促进材料发现和开发 [14]。会上，材料基因组计划小组委员会琳达・萨博卡（Linda Sapochak）和恰克・沃德（Chuck Ward）做了题为《为未来十年材料科学的加速发展做好准备》的主题演讲。会议围绕 5 个主题展开交流和讨论：一是如何加强学术界、政府和联邦实验室的互动；二是研究中心和研究平台如何为材料基因组计划相关研究提供焦点；三是材料基因组计划如何应对贯穿材料领域的主要挑战；四是数据驱动的研究如何实现技术材料发现和开发；五是材料发现如何面向开发。

2　人工智能在材料基因组技术中的应用成为热点

美国材料基因组计划通过高通量计算和高通量实验来加速材料的发现周期，可以用来解决今天面临的技术和社会挑战。虽然材料基因组计划的努力已经成功地在数千种材料中筛选出一些有用的材料，但在元素周期表中有 100 多种元素，这些元素可组合形成大量潜在的新材料，组合可能性超过了 10^{100} 种，如果没有指导，在这个巨大的组合空间中进行搜索将会是缓慢、昂贵并令人沮丧的 [15]，也会严重限制材料基因组计划理念的进一步实施 [16]。

解决筛选新材料海量组合问题的方法之一是

使用人工智能工具，例如利用机器学习、深度学习等人工智能优化技术，来有效地评估海量的材料组合，从而达到进行功能预测和材料设计的目的。虽然人工智能在语音识别、汽车自动驾驶、语言翻译等领域非常成功，但在材料领域还处于初级发展阶段[16]。

近来，人工智能的应用逐渐成为美国材料基因组技术的研究热点。关于人工智能和材料基因组技术的学术论文数量几乎呈指数级增长[17]，文献检索中“材料信息学”“机器学习”都成为材料研究的热门话题[5]。美国材料基因组计划负责人、科学家相继发表文章或演讲，对人工智能在材料基因组技术中的应用和发展进行深入探讨。2017年10月麻省理工学院材料科学家艾尔莎·奥利维蒂（Elsa Olivetti）等发表论文《通过文本提取和机器学习从科学文献中获取材料的“配方”》，研究使用人工智能和机器学习技术从数以百万计的材料研究论文中提取生产材料的“配方”[18]。2019年4月西北大学材料科学家克里斯托夫·沃尔弗顿（Christopher Wolverton）、斯坦福大学SLAC国家加速器实验室科学家阿佩瓦·梅赫塔（Apurva Mehta）等发表论文《通过迭代机器学习和高通量实验，加速金属玻璃的研发》，研究用机器学习模型与高通量实验相结合的方法，来指导金属玻璃材料的设计开发。2019年6月美国国家科技委员会材料基因组计划执行秘书、美国标准技术研究所材料测量实验室材料基因组学技术项目主任詹姆斯·A.沃伦发表论文《材料基因组计划与人工智能》，研究人工智能在材料基因组技术中的应用情况和发展方向。

美国标准技术研究所、美国能源部、材料研究学会等先后组织多场关于人工智能和材料基因组技术的研讨会。2017年8月，美国能源部先进制造办公室联合工业界、学术界组织召开了“人工智能应用于材料发现和设计研讨会”（Workshop on Artificial Intelligence Applied to Materials Discovery and Design），汇聚了100多位杰出的科技专家，共同探讨人工智能应用于材料研究和开发的机遇和挑战，推动开发基于人工智能的数字化工具，推动人工智能最新技术在能源等相关领域材料的设计和发现中得到应用，以发现和设计可以提高美国制造业竞争力的创新材料。研讨会围绕数据量和数据质量、材料创新平台和基础设施、人工智能在材料设计中的特定应用3个主题，主要探讨了材料研发中的一些共性问题，例如数据格式、算法、模型和工具，一些研究人员分享了将人工智能技术用于合金设计制造以及玻璃和催化剂开发方面的经验[19]。2017年11月，材料研究学会举行秋季大会，美国国家科技委员会材料基因组计划执行秘书、美国标准技术研究所材料测量实验室材料基因组学技术项目主任詹姆斯·A.沃伦参加并发表了《材料基因组计划与人工智能》主题演讲[20, 21]。2018年8月7—8日，美国标准技术研究所组织召开了“材料科学与人工智能研讨会”（Workshop Artificial Intelligence for Materials Science），针对人工智能技术应用于新材料研发面临的两个挑战进行研讨[16]。研讨会中不少演讲主题围绕人工智能，特别是机器学习在材料研发和设计中的应用，会上，詹姆斯·A.沃伦还发表了《材料基因组计划与人工智能》的主题演讲[16]。

这些论文、会议和演讲对人工智在材料基因组技术中的应用情况和发展前景进行了深入研究和讨论，提出了很多新观点。

2.1 人工智能在材料基因组技术中的应用取得进展

人工智能被越来越多地应用在材料科学领域，以期发现运用传统手段无法发现的现象和规律。很多参与材料基因组计划的科研团队通过引入人工智能技术，例如机器学习、深度学习等算法来分析海量的实验结果，预测新材料的特性，指导设计新材料，取得了很多进展。

一是2018年6月詹姆斯·沃伦发表的论文《材料基因组计划与人工智能》显示，美国国家标准与技术研究院围绕数据基础设施需求建立了对材料基因组计划的强大支持，该数据基础设施能够快速发现现有材料数据和模型，评估和改善这些数据质量，最后根据这些数据开发新的方法和计量工具。通过政府、学术界与工业界机构的紧密合作，这些方法正在取得重大进展[5]。

二是西北大学材料科学家克里斯托夫·沃尔弗顿、斯坦福大学SLAC国家加速器实验室科学家阿佩瓦·梅赫塔的研究团队使用人工智能来研究制造新型金属玻璃，进展速度比在实验室快

200 倍[22]，并期望该方法未来可以用于设计手机和宇宙飞船的构成材料的开发。沃尔弗顿称，科学家们可以同时合成和测试数千种材料，但即使以这样的速度，盲目地尝试各种可能的组合也是在浪费时间。人工智能可以大幅减少实验数量，科学家告诉人工智能他们想要创建材料的特性，人工智能会为科学家筛选出备选的实验组合来制作新材料。

三是麻省理工学院材料科学家艾尔莎·奥利维蒂、马萨诸塞大学阿默斯特分校和加州大学伯克利分校的研究团队开发了一种机器学习系统，可以扫描材料科学论文，自动搜索和提取材料科学论文中各种材料的“配方”[23]。该机器学习系统现在可以以 99% 的准确度检测到哪些段落包含材料“配方”，以 86% 的准确度找到段落中正确的词语[22]。该机器学习系统还可以分析提取数据，以推断材料类别的一般特征，例如合成材料所需的不同温度范围，以及材料的特殊特征[23]。

四是卡耐基·梅隆大学布莱恩·迪克斯特（Brian DeCost）博士及其同事将机器学习技术应用于合金粉末增材制造。研究人员以粉末的显微照片为起点，通过定义的描述符（局部几何的各种短长度尺度描述），使用来自计算机视觉领域的方法来检测这些特征描述符，并将特定合金模式与特定特征相关联。这种方式类似于人们分析微观结构，但是这样的算法可以找到人类不太可能识别的相关性。有时相关性是虚假的，但是足够的算法训练和交叉验证提高了这些方法的可靠性。目前研究人员能够以 95% 的准确率识别金属粉末类型，这比人类的表现更好。结构的表征是材料科学的核心，人工智能能够提供有力帮助[5]。

五是斯坦福大学材料科学与工程系助理教授伊万·瑞德（Evan Reed）领导的团队一直在使用人工智能辅助开发更好的锂离子电池电解质。通常电解质由多种材料组成，寻找材料的最佳组合是非常困难的。瑞德称，他们已经开发出一种机器学习模型，在预测使用哪种材料方面，其表现超过了专家的经验和直觉[5]。

六是国家标准技术研究所吉拉德·库什尼（Gilad Kusne）、SLAC 国家加速器实验室和马里兰大学的合作团队使用人工智能快速测量绘制三元相图。利用机器学习对 X 射线束的光栅扫描图案进行图像识别和分析测定，将绘制出的相图通过抽样实验进行验证。这种高通量的实验与人工智能相结合的方式，可以快速高效地来分析完成任务[5]。

2.2　机器学习、深度学习等算法是应用的关键

将人工智能技术方法应用于材料发现和设计是一项涉及多个学科的工作，需要材料科学、化学、数学和计算机科学等许多学科的知识。材料基因组计划在努力开发能够创建和捕获材料知识的机器学习系统，并使系统和知识能为普通的材料研究人员使用。因此必须构建人工智能系统模型，以紧凑的形式封装大量有关材料科学的信息，如同构建一个黑匣子，使得研究人员能够将微尺度现象与宏观性质关联起来，而不必去考虑中间尺度。人工智能就是创建模型的一种方法。但如何找到数据驱动方法和传统物理实验模型的结合点，取决于机器学习、深度学习等算法以及描述符（输入）和预测（输出）。尽管存在很多挑战，难度很大，但材料领域的很多科学家对机器学习、深度学习、神经网络（Neural Network）、优化算法（Optimization Algorithm）等在材料科学领域应用的潜力和前景都抱有很高的期望。

2019 年 3 月詹姆斯·A. 沃伦在接受 Singularity Hub 网站采访时表示，机器学习是未来材料科学发展的关键，将帮助科学家打破目前材料的理论极限，有可能会帮助形成许多令人兴奋的新材料的发现[22]。沃伦称，机器学习和人工智能应用程序在材料数据库的再次加工处理中应用潜力很大，可以形成很多发现新材料的机会[5]。

斯坦福大学 SLAC 国家加速器实验室科学家阿佩瓦·梅赫塔称，他们使用人工智能来研究新金属玻璃材料的方法类似于人们学习新语言的方式（坐下来记住所有语法规则，通过倾听别人的谈话积累经验）。首先，研究人员通过发表的论文查看了尽可能多的关于如何制作不同类型金属玻璃的数据，然后将这些“语法规则”输入到机器学习算法中。最后，算法学会自己预测哪些元素组合会创造一种新形式的金属玻璃。梅赫塔的研究团队随后在实验室中验证了该方法。此外，进行实验意味着他们现在有更多的数据可以反馈

给算法，因此算法每次都会变得更聪明，更智能[22]。

麻省理工学院材料科学家艾尔莎·奥利维蒂、马萨诸塞大学阿默斯特分校和加州大学伯克利分校的研究团队使用有监督和无监督的机器学习技术对他们的机器学习系统进行训练。由于材料配方提取是一个新的研究领域，奥利维蒂和她的同事们并没有多年来由不同研究团队积累的大量注释数据集。他们必须自己注释他们的 100 多篇论文数据，但这对于机器学习标准是一个非常小的数据集。为了改进机器学习系统，他们使用了谷歌开发的一种名为 Word2vec 的算法。Word2vec 查看单词出现的上下文，并将那些倾向于具有相似上下文的单词组合在一起。使用 Word2vec，研究人员能够极大地扩展他们的训练集，因为机器学习系统可以推断附加到任何给定单词的标签可能适用于与其聚集的其他单词。因此，研究人员可以在大约 64 万篇论文上训练他们的系统，远远超过 100 篇的规模[23]。研究人员还在探索一系列深度学习技术，这些技术可以进一步概括材料配方的结构，目的是自动设计材料配方中未考虑到的材料，他们希望进一步的工作能够提高系统的准确性和完整性[23]。

2.3 人工智能将推动材料基因组技术加速发展

数百年来，人们都是通过反复试验或运气和偶然性发现新材料的。现在，科学家正在使用人工智能来加速这一过程，并可以在进入实验室之前提前预测哪些化合物可以设计出想要的材料。随着人工智能迅速崛起，人工智能在材料基因组计划、材料设计方面发挥了重要作用，该领域的研究工作出现了爆炸式的增长[5]。大数据分析技术与各领域深度融合，数据挖掘技术和机器学习算法被应用到材料学领域，人工智能正在深刻影响材料研究领域。The Verge 网站在 4 月 25 日采访了詹姆斯·A. 沃伦和克里斯托夫·沃尔弗顿、阿佩瓦·梅赫塔、艾尔莎·奥利维蒂等知名材料科学家，发表了《人工智能如何帮助我们以前所未有的速度发现材料（How AI is helping us discover materials faster than ever）》。沃伦和材料基因组计划的科学家们也纷纷预测，人工智能将推动材料基因组技术加速发展。

2019 年 4 月克里斯托夫·沃尔弗顿、阿佩瓦·梅赫塔等发表的论文《通过迭代机器学习和高通量实验，加速金属玻璃的研发》显示，在寻找 Co–V–Zr 中新的金属玻璃的过程中，机器学习与高通量实验结合有助于加速材料的发现速度（见图 3）。实验观察结果与模型预测结果非常一致，但预测的成分存在定量差异。他们使用这些差异再来重新训练机器学习模型，模型显著提高了准确度。通过迭代使用机器学习和高通量实验，他们快速地发现了 3 种新的金属玻璃，有望大大加速新的金属玻璃材料开发过程。他们相信这种开发范式也可以适用于更广泛的材料研究[15]。

马里兰大学瓦伦汀·斯丹尼（Valentin Stanev）教授在超导体材料研究中引入机器学习。斯丹尼表示，他们列出了所有已知的超导体材料，并将机器学习应用到开发流程中，帮助找到新的超导体材料。斯丹尼称，机器学习在科学过程中的新应用可以将运行实验所需的时间减少高达 80%。在实验室建立机器学习系统，运用机器学习算法就能根据实验结果决定接下来做什么实验，并预测实验的总体结果。从某种意义上说，可能只需要运行 10%~20% 的实验来获得 100% 的图像。还可以将实验的部分控制权交给人工智能系统，由人工智能系统自主决定下一步要采取什么措施[17]。

斯坦福大学伊万·瑞德教授称，机器学习甚至可能用于某种“逆向工程”：假如你需要一种具有某种特性的电池，你可以将这些信息输入到机器学习模型中，然后自动运行所有可用的已知材料，机器学习模型就可以提出一系列可能符合要求的由不同材料组成的电池。除了超导体和热电材料之外，科学家们认为机器学习可能会促进燃料电池储氢装置的进步。在医疗保健方面，它可以帮助设计能更好地控制药物溶解时间的新材料[17]。

詹姆斯·A. 沃伦表示，在过去 5 年中，使用人工智能和机器学习技术的材料科学研究团队发展迅速，取得爆炸式的发展。关于这一主题的科学论文数量几乎呈指数级增长，我们已经看到了以研究为基础的现实世界的进步，但这还仅仅只是开始。机器学习可以使新材料的科学开发和改进过程的每

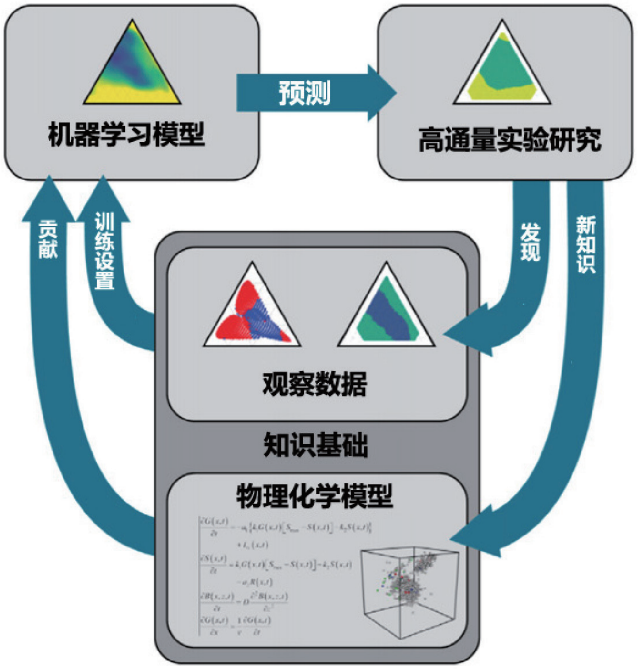

图 3　机器学习与高通量实验结合将加速材料发现 [15]

一步都受益 [22]。詹姆斯·A. 沃伦认为，意识到我们可以通过在“最佳”位置进行测量来增加知识后，我们也可以让人工智能负责制作样本。我们可以合成人工智能指示的样本，以提供增加知识的最大可能性，然后测量该样本并重复。由于人工智能不需要休息，人们可以使用更多的组件，并允许处理技术的大变化 [21]。詹姆斯·A. 沃伦认为，人工智能驱动的材料基因组技术与制造方法的持续普及，对社会和经济的影响将是非常深刻的 [5]。

2.4　人工智能在材料基因组技术中应用仍面临不少挑战

詹姆斯·A. 沃伦和材料基因组计划的科学家们认为，人工智能和材料科学结合的未来很有希望，但挑战仍然存在 [22]。

一是设计人工智能时如何选择材料的有效描述符（即有效的输入）。输入信息的不足会直接导致人工智能模型无法准确预测新材料的性能。瑞士联邦理工学院洛桑分校的尼古拉·马扎里（Nicola Marzari）团队开发了一种人工智能算法来筛选来自多个数据库的信息，查找可以剥离的 3D 材料，以创建仅一层的 2D 材料，如石墨烯材料。尼古拉·马扎里称，由于常常采用真实材料的简化模型，而各种环境因素变化都会影响化合物的性能，因此人工智能预测结果常常会存在偏差，大多数人工智能模型都不能将这些因素考虑在内。所以他仍不得不通过自己的经验告诉人工智能算法应当遵循某些规则，比如去寻找弱化学键 [22]。根据尼古拉·马扎里 2019 年 3 月在《自然纳米技术》上发表的论文，团队开发的人工智能算法从超过 10 万种材料中发现大约 2 000 种材料可以剥离成 2D 材料。尼古拉·马扎里称，这些材料是一个“宝库”，因为许多材料可以改善电子产品的特性 [22]。

二是设计人工智能时如何选择合适的算法和工作流程，如在人工智能模型中如何引入基于物理学的材料模型 [16]。詹姆斯·A. 沃伦称，机器学习、深度学习和神经网络算法的挑战在于复合材料的数学模型本质上是高度多维的，因为材料属性具有如此多的变量，但这也正是人工智能未来可以贡献力量的地方。斯托夫·沃尔弗顿称，希望有一天，只需要人工智能和机器人在实验室研究开发材料，而根本不需要人类去做实验，希望未来可以创建一个没有任何人参与的、真正完全自治的智能系统 [22]。

三是目前人工智能没有足够大的化合物性能

数据库支撑。斯托夫·沃尔弗顿表示，缺乏数据意味着人工智能算法不是很聪明[22]。詹姆斯·A. 沃伦称，材料数据是材料基因组计划的“燃料”，但目前获取这些信息的方法非常有限，通常是通过高通量实验和挖掘已公开发布的数据。不幸的是很多实验结果并不公开发表，也就无从获取数据。这就需要搭建材料数据的共享平台和建立激励措施来鼓励研究人员公开发布更多数据[21]。伯克利实验室计算化学博士后研究员夏姆·德怀克纳（Shyam Dwaraknath）认为，我们刚刚进入材料科学领域的大数据时代，拥有了不少精心策划和直接可比数据的大型数据库，但材料的真正复杂性远远大于此，现在互联网上的所有数据字节的数量才刚刚相当于一粒沙子中的原子数量[22]。

3 相关建议

我国材料基因组研究最早开始于2011年12月召开的香山科学会议，会议当时提出尽快建立以高通量材料计算模拟、高通量复合材料实验、材料共享数据库为基础的“材料基因组计划”平台。2012年12月，由中国工程院领衔的“材料科学系统工程发展战略研究——中国版材料基因组计划”重大项目启动。2016年2月，科技部启动了“材料基因工程关键技术与支撑平台”重点专项，共部署40个重点研究任务，实施周期为5年，自此材料基因组研究进入了一个快速发展时期。缩短新材料的“发现—研发—生产—应用”周期，降低材料研发中的人力、物力成本是实施材料基因组计划的根本目的，也是中国实现新材料领域跨越式发展的内在需求[24]。

美国材料基因组计划起步早，但在实施过程中遭遇发展瓶颈，进展缓慢。近两年随着人工智能技术的飞速发展，美国大数据技术、机器学习技术等在材料科学领域得到广泛应用，人工智能技术在材料基因组研究中的应用已经成为研究热点，取得了不少进展，未来材料基因组技术可能加速发展。我国应密切关注，提早布局。

（1）加快人工智能在材料基因组技术中的应用。

人工智能在辅助解决计算平台、试验平台、数据平台三大平台的瓶颈问题，提高新材料开发效率方面有非常显著的作用。近年来我国人工智能技术发展迅速，应加快人工智能技术在材料基因组技术中的应用，加大对人工智能算法、人工智能模型引入基于物理学的材料模型等关键技术的研究，争取尽早取得技术突破，从而提高新材料的研发效率，提升应对高性能新材料需求的快速反应和生产能力。

（2）加强材料数据共享平台建设，支撑材料基因组技术加速发展。

数据是材料基因组计划的“燃料”，如同生物基因组计划一样，材料基因组技术所需数据量庞大，我国发展人工智能具有数据量、样本量大的天然优势。因此加强材料数据共享平台建设，统一数据标准接口，有利于集合各方面优势资源，凝聚人力物力，形成规模庞大的材料基础数据库，为人工智能在材料基因组技术中的应用提供充足“动能”。

（3）密切跟踪美国材料基因组计划进展，加强材料基础研究国际合作。

美国是最早设立材料基因组计划的国家，在目标规划、组织协调、研究机制等方面还有不少先进的理念和做法值得密切跟踪，并结合国情加以学习借鉴。同时，材料基因组计划任务量巨大，需要各国通力合作进行开发合作，汇集优势资源共同攻关，我们应当进一步加强材料基础研究的国际合作。■

参考文献：

[1] 杨小渝，王娟，任杰，等. 支撑材料基因工程的高通量材料集成计算平台 [J]. 计算物理，34（6）：697-704.

[2] 郭艳华，刘畅. 走进微观世界，探秘材料基因——材料基因组论坛侧记 [J]. 中国材料进展，2017（11）：798-799.

[3] 林海，郑家新，林原，等. 材料基因组技术在新能源材料领域应用进展 [J]. 储能科学与技术，2017，6（5）：990-999.

[4] National Science and Technology Council, Office of Science and Technology Policy. Materials Genome Initiative for global competitiveness[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/materials_genome_initiative-final.pdf.

[5] Warren A James. Materials Genome Initiative and

artificial intelligence[EB/OL]. [2019-06-11]. https://doi.org/10.1557/mrs.2018.122.

[6] National Science and Technology Council, Office of Science and Technology Policy, Subcommittee on the Materials Genome Initiative. Materials Genome Initiative Strategic Plan[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/mgi_strategic_plan_-_dec_2014.pdf.

[7] Executive Office of the President. Fact Sheet: Progress on Materials Genome Initiative[EB/OL]. [2019-06-11]. https://www.whitehouse.gov/.

[8] Executive Office of the President. Fact Sheet: Materials Innovation-Materials Genome Initiative turns two[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/2013_mgi_annoucements.pdf.

[9] Office of Science and Technology Policy. Fact Sheet: The Materials Genome Initiative-Three years of progress[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/materials_genome_initiative_-_three_years.pdf.

[10] Thomas Kalil, Lloyd Whitman. The Materials Genome Initiative: The first five years[EB/OL]. [2019-06-11]. https://obamawhitehouse.archives.gov/blog/2016/08/01/materials-genome-initiative-first-five-years.

[11] Office of Science and Technology Policy. The first five years of the Materials Genome Initiative: Accomplishments and technical highlights[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/mgi-accomplishments-at-5-years-august-2016.pdf.

[12] US Department of Energy. Materials Genome Initiative 2nd Annual Principal Investigator Meeting[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/2015_MGI_PI_Meeting_Abstract_Book.pdf.

[13] MGI. Materials Genome Initiative 3rd Annual Principal Investigator Meeting[EB/OL]. [2019-06-11]. https://www.mgi.gov/sites/default/files/documents/2016_Abstract_Book_Final.pdf .

[14] National Science Foundation. Materials Genome Initiative 4th Annual Principal Investigator Meeting[EB/OL]. [2019-06-11]. https://www.nsf.gov/attachments/297691/public/2018_MGI_PI_Meeting_Abstract_Book_Draft.pdf.

[15] Ren F, Ward L, Williams T, et al. Accelerated discovery of metallic glasses through iteration of machine learning and high-throughput experiments[J]. Science Advances, 2018, 4(4): 1 566.

[16] National Institute of Standards and Technology. Workshop Artificial Intelligence for Materials Science (AIMS) [EB/OL]. [2019-06-11]. https://www.nist.gov/news-events/events/2018/08/workshop-artificial-intelligence-materials-science-aims.

[17] Marc Prosser. The Wild New Materials of the future will be discovered with AI[EB/OL]. [2019-06-11]. https://singularityhub.com/2018/03/22/how-the-wild-new-materials-of-the-future-will-be-discovered-with-ai/.

[18] Edward Kim, Kevin Huang, Adam Saunders, et al. Materials synthesis insights from scientific literature via text extraction and machine learning[J]. Chemistry of Materials, 2017, 29 (21): 9 436-9 444.

[19] US Department of Energy. Workshop on artificial intelligence applied to materials discovery and design[EB/OL]. [2019-06-11]. https://www.energy.gov/eere/amo/downloads/workshop-artificial-intelligence-applied-materials-discovery-and-design.

[20] Materials Research Society. Plenary session featuring the Fred Kavli distinguished lectureship in materials science[EB/OL]. [2019-06-11]. https://www.mrs.org/fall-2017-plenary-session.

[21] Materials Research Society. Plenary session featuring the Fred Kavli distinguished lectureship in materials science. [2019-06-11]. https://materials.typepad.com/mrs_meeting_scene/2017/11/plenary-session-featuring-the-fred-kavli-distinguished-lectureship-in-materials-science.html.

[22] Angela Chen. How AI is helping us discover materials faster than ever[EB/OL]. [2019-06-11]. https://www.theverge.com/2018/4/25/17275270/artificial-intelligence-materials-science-computation.

[23] Larry Hardesty. Artificial intelligence aids materials fabrication[EB/OL]. [2019-06-11]. http://news.mit.edu/2017/artificial-intelligence-aids-materials-fabrication-1106

[24] 沈自才，代巍，马子良 . 航天材料基因工程及其若干关键技术 [J]. 航天器环境工程，2017，34（3）：324-329.

Artificial Intelligence will Accelerate the Development of Material Genome Technology

HUANG He, CHEN Hong-sheng
(Ministry of Science and Technology of the People´s Republic of China, Beijing 100862)

Abstract: Through the combing and analysis of the introduction and implementation of the US Material Genome Initiative, this paper finds that the application of artificial intelligence in the US material genome technology has become a research highlight, among which algorithms such as machine learning and deep learning are the key. Artificial intelligence is very effective in solving the bottleneck problem of the Material Genome Initiative, and it may promote the accelerated development of material genome technology, but it also faces many challenges. On this basis, this paper proposes relevant suggestions for accelerating the research of material genomics technology in China.

Key words: US Material Genome Initiative; artificial intelligence; machine learning; deep learning

(上接第26页)

The Evaluation of China's National Competitiveness and Comparison between China and the U.S.

HAN Jia-wei, XUAN Zhao-hui
(Chinese Academy of Science and Technology for Development, Beijing 100038)

Abstract: The Global Competitiveness Report is one of the most influential reports that evaluate economies' competitiveness. After introducing the history and indicator matrix of the report, this paper analyzes the strength and weakness of China's competitiveness and compares that between China and the U.S. on the basis of newly published Global Competitiveness Report 2019. The results show that the gap of competitiveness between China and leading economies is narrowing. China's strength lies in macroeconomic stability, financial system, infrastructure and innovation ecosystem, while weakness lies in markets, human capital and institutions. China has to make full use of the advantages and make up for weakness to improve its international competitiveness and economic growth potentials.

Key words: The Global Competitiveness Report; competitiveness; indicator matrix; comparison between China and the U.S.