

人工智能在材料科学领域的应用发展研究 ——基于 WOS 核心数据库的文献计量分析

褚旭龙^{1,2}, 刘进长¹, 史冬梅¹

(1. 科学技术部高技术研究中心, 北京 100044;

2. 北京航空航天大学公共管理学院, 北京 100191)

摘要:以实验试错法和基于密度泛函理论的方法为代表的传统新材料发现方法, 逐渐在开发成本、效率以及周期等方面显现出不足。人工智能, 特别是机器学习和深度学习, 综合强大的数据处理和高预测性能, 在材料发现、材料设计、材料检测和材料分析等领域展示出优势, 对新材料研发创新具有重要推动作用。以近 10 年美国信息情报研究所 (ISI) 下属的 Web of Science (WOS) 核心数据库收录的相关文献为数据源, 使用 CiteSpace 软件开展可视化分析, 旨在总结分析人工智能在材料科学领域发展的研究现状、研究热点及其研究发展趋势, 为中国人工智能与材料科学融合创新发展提供相关参考, 以期提升研究水平。

关键词:人工智能; 交叉科学; 材料科学; 文献计量

中图分类号: G321 **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2023.04.008

材料作为高新技术应用发展的基础, 是先进制造产业发展的先导与保障, 对加速技术融合创新与产业优化提升具有极其重要的驱动作用, 是事关国家经济、社会、军事安全发展的重要科技领域^[1]。近年来, 主要发达国家针对关键材料领域纷纷出台相关政策, 积极开展战略布局, 精细化项目扶持, 以期提升自身材料产业化能力。同时, 针对重要材料领域制定了维护本国利益的贸易政策, 对关键技术或核心产品设置出口壁垒, 如美国商务部工业与安全局提出的《出口管制框架方案》中列出了 14 个管制领域, 涉及先进材料技术及其相关的制造设备与检测仪器等^[2]。如今, 中国新型工业化步伐显著加快, 如何推动材料领域的创新发展尤为关键, 具有重要战略意义。此外, 以智能制造技术、新一代信息技术以及新能源技术等为代表的新兴产业快速发展, 对材料提出了更高的要求, 包括高性能、低

缺陷、高速迭代、高耐用、低成本和易回收等, 新材料的研发面临空前挑战^[3]。

人工智能, 特别是机器学习和深度学习算法, 能够预测材料性能, 设计和发现新材料, 降低研发成本并缩短研发周期, 成为材料科学及新材料科学领域研究发展的重要驱动力。人工智能能够在材料组成元素特征、原子结构等信息与预期目标属性之间建立预测关系, 并在材料的带隙、相稳定性、电导率、导热系数、熔融温度等性能预测方面取得了很好的应用效果^[4]。因此, 加速推进人工智能与新材料技术的融合创新, 对提升中国在材料科学领域的科技创新水平起着至关重要的作用。

综上所述, 本文开展国内外人工智能在材料科学领域应用发展现状、趋势以及前沿热点研究。以 2012—2022 年美国信息情报研究所 (ISI) 下属的 Web of Science (WOS) 核心数据库文献为数据源, 选用

第一作者简介: 褚旭龙 (1988—), 男, 博士, 助理研究员, 主要研究方向为高技术领域前沿热点研究。

通信作者简介: 史冬梅 (1964—), 女, 博士, 研究员, 主要研究方向为科技发展战略与科技计划管理。电子邮箱: shidm@htrdc.com

收稿日期: 2023-02-20

CiteSpace 软件开展可视化分析, 并对生成的可视化结果进行归纳与总结, 旨在为中国人工智能与材料科学融合创新发展提供相关参考, 以期提升研究水平。

1 数据源与分析方法

1.1 数据源

本文选取 WOS 核心数据库为文献来源, 为保证研究数据全面性和高质量性, 分别在检索策略和文献筛选等环节进行了研究设计。在检索策略设计环节, 依据研究主题分别从人工智能技术和材料应用两部分逐一开展设定。在人工智能技术关键词设定方面, 本文按照其算法的分类关系进行检索关键词的分解设定。首先, 人工智能的主流算法为“机器学习”, 机器学习可分为“监督学习”“无监督学习”以及“半监督学习”等类型(监督学习又可分为“回归”与“分类”), 其常用算法具体包括“神经网络”“支持向量机”“决策树”“随机森林”“强化学习”等; 另一主流算法为“深度学习”, 又可分为“卷积神经网络”“循环神经网络”等; 另外, “遗传算法”也是人工智能应用于材料研发的热点算法。在新材料技术关键词设定方面, 本文按照材料应用的前后阶段, 设定检索关键词为“材料发现”“材料设计”“材料检测”和“材料分析”。考虑同义词和缩写等情况并按照逻辑“与”的关系进行组合, 经过多次尝试, 最终本文的检索策略设定为 TS=("artificial intelligence" OR "AI" OR "machine learning" OR "ML" OR "deep learning" OR "DL" OR "neural network" OR "reinforcement learning" OR "genetic algorithm" OR "support vector machine" OR "SVM" "decision tree" OR "random forest" OR "unsupervised learning") AND TS=("material discovery" OR "material design" OR "material testing" OR "material analysis")。在该检索策略下, 将检索方式设定为主题检索, 期刊发表年份范围设置为“2012—2022 年”, 共计检索出相关期刊文献 506 篇。导出所有检索记录(作者、标题、摘要、关键词和引用的参考文献等), 并将其转化为 CiteSpace 可处理的数据格式。

在文献筛选环节, 为保证研究数据的高质量, 对检索出的文献进行了多重筛选。先对所检索出的文献进行逐一核查, 剔除了会议通知、书稿序言、

报纸等研究价值较低的文献, 以及与本文研究主题相关度不高的文献, 并对剩余的文献数据进行去重操作。通过上述步骤, 最终得到了本文的研究数据, 共计 431 篇高质量文献。

1.2 分析方法

本文选用 CiteSpace 软件, 以知识图谱的方式直观展示目标研究领域内在相关主题间的逻辑关系, 进而开展研究现状、研究热点的分析^[5]。研究要素方面, 分别从国家、机构、关键词和被引文献等不同方面开展分析, 对应 CiteSpace 软件设置为“Country”“Institution”“Keyword”“Reference”。聚类分析采用 LLR (Log Likelihood Ratio) 算法计算和确定每种标签的类型。

最终, 本文对纳入的 431 篇高质量文献从时间分布、空间分布、研究机构分布、关键节点核心文献、关键词共现、关键词突现及时间线分析等方面进行梳理, 梳理人工智能在材料科学领域的研究现状、发展脉络、研究热点和新兴热点, 总结该领域的关键研究内容, 揭示其未来发展趋势。

2 结果与分析

2.1 时间分布特征

时间分布特征主要依据年发文量开展分析, 可以在一定程度上了解人工智能在材料科学领域应用发展的总体趋势。如图 1 所示, 从 2012 年 1 月至 2022 年 10 月, 有关人工智能在材料科学领域应用的文献发文量整体呈现稳中有升的发展趋势。且以 2018 年为分界点, 该领域的相关研究可分为两个阶段。第一个阶段为 2018 年以前, 该领域的年度平均发文量约为 9 篇。第二阶段为 2018 年及以后, 该领域的年度发文量约为 76 篇, 且逐年发文上升趋势幅度大。

2016 年被称为“人工智能元年”^[6], 人工智能得到了技术突破与广泛应用, 并推动了相关理论和共性技术的发展。其中, 材料科学领域作为人工智能重要的应用领域, 逐步受到研究者的关注。总体而言, 人工智能在材料科学领域应用的研究成果体现出持续快速增长特征。

2.2 空间分布特征

空间分布特征主要依据国家合作关系开展分析。国家合作网络图谱能够一定程度地了解研究领

域的地理分布和合作情况。如图 2 所示，图谱中国国家节点数为 31 个，合作关系连线为 49 条，网络整体密度为 0.112 6。

图中节点和字体的大小代表国家发表论文数量的多少，可以看出，中国在该领域研究中发文量排名第一位，为 166 篇；美国发文量排名第二位，

为 121 篇；日本发文量排名第三位，为 36 篇。同时，表 1 列出了各个国家在该领域发文的具体篇数（频次）和中心性取值。研究主体的中心性越高，就代表与其他研究主体之间的联系越紧密，在领域中的影响力越大。中心性取值排名前三位的国家分别是中国、美国和德国。

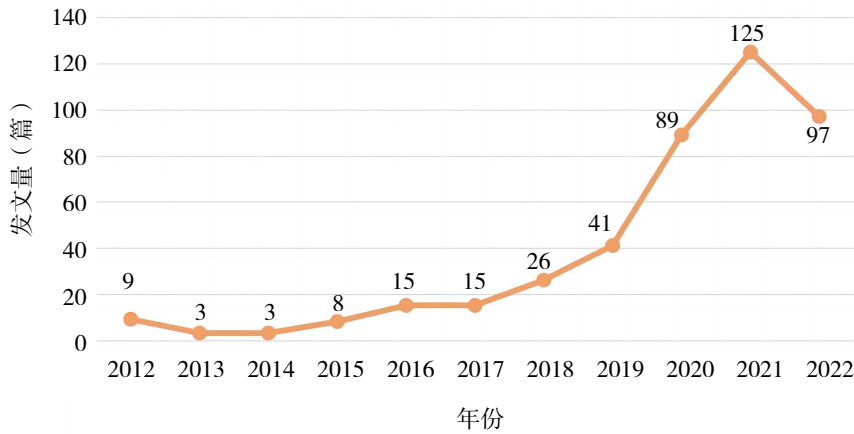


图 1 2012—2022 年 10 月人工智能在材料科学领域应用的文献发文量

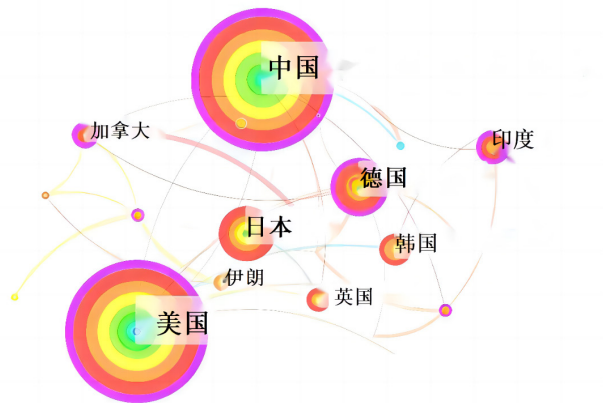


图 2 国家合作关系分布

表 1 2012—2022 年 10 月国家发文量排名前 10 位的分布情况

序号	国别	篇数（频次）	中心性
1	中国	166	0.59
2	美国	121	0.49
3	日本	36	0.03
4	德国	32	0.34
5	韩国	26	0.09

续表

序号	国别	篇数（频次）	中心性
6	印度	18	0.24
7	伊朗	14	0.09
8	英国	13	0.00
9	加拿大	12	0.28
10	波兰	5	0.00

2.3 研究机构分布特征

研究机构分布特征主要依据研究机构的发文情况开展分析，本文对纳入的 431 篇高质量文献进行研究机构发文量统计分析，从而了解人工智能在材料科学领域应用发展研究的关键研究机构。如图 3 所示，可以看出，中国科学院在该领域的发文量排名第一位，为 14 篇；麻省理工学院和清华大学发文量排名第二位，均为 10 篇；哈尔滨工业大学和北京科技大学排名第三位，均为 8 篇。

同时，本文使用 CiteSpace 软件进行研究机构合作网络图谱分析，旨在展示出研究领域的核心研究

群体。如图4所示,该领域形成了9个合作性较强的研究机构群体。规模最大的集群涉及19个节点,主要包括15家中国研究机构(中国科学院、清华大学、哈尔滨工业大学以及浙江大学等)、4家美国研究机构(美国空军大学、约翰霍普金斯大学、普渡大学、内布拉斯加大学)和1家德国研究机构(亚琛工业大学)。该集群主要以中国科学院、清华大学、哈尔滨工业大学、北京科技大学等研究机构为关键节点形成合作网络关系,在一定程度上反映出中国在人工智能与材料科学融合发展领域具备突出的研究活跃度和产出量。该集群内跨国合作关系具体为哈尔滨工业大学与约翰霍普金斯大学/美国空军大学、北京科技大学与普渡大学/内布拉斯加大学、亚琛工业大学分别与北京科技大学和中国科学院的合作等。

其他跨国合作关系群体主要包括美国斯坦福大学和加拿大多伦多大学、美国爱荷华州立大学和

日本东京大学、美国加利福尼亚大学伯克利分校和韩国科学技术院等,可以看出美国在国际合作方面具备较大优势。全球发文量排名第二位的麻省理工学院主要研究作为独立研究,并未与国内或国外研究机构合作,但其研究成果数量同样突出。

2.4 关键节点核心文献分析

关键节点核心文献是基于文献被引频次分析程序统计得出的重要参考文献,具有高被引、高中心度等特征,能够反映出研究者所关注的焦点和未来的研究方向^[7]。如表2所示,对纳入的431篇高质量文献进行人工智能在材料科学领域的关键节点核心文献被引频次统计,按照领域被引频次关系,梳理了2012—2022年10月WOS核心数据库中人工智能在材料科学领域应用发展相关研究排名前10位的关键节点核心文献,同时也列出了对应文献的网络总被引频次。

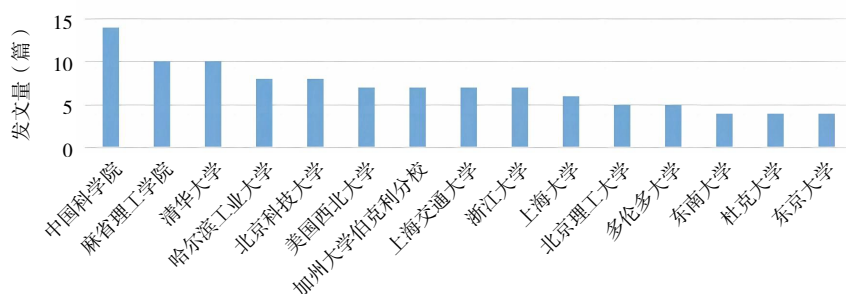


图3 2012—2022年10月人工智能在材料科学领域应用发展研究相关发文量排名前15位研究机构

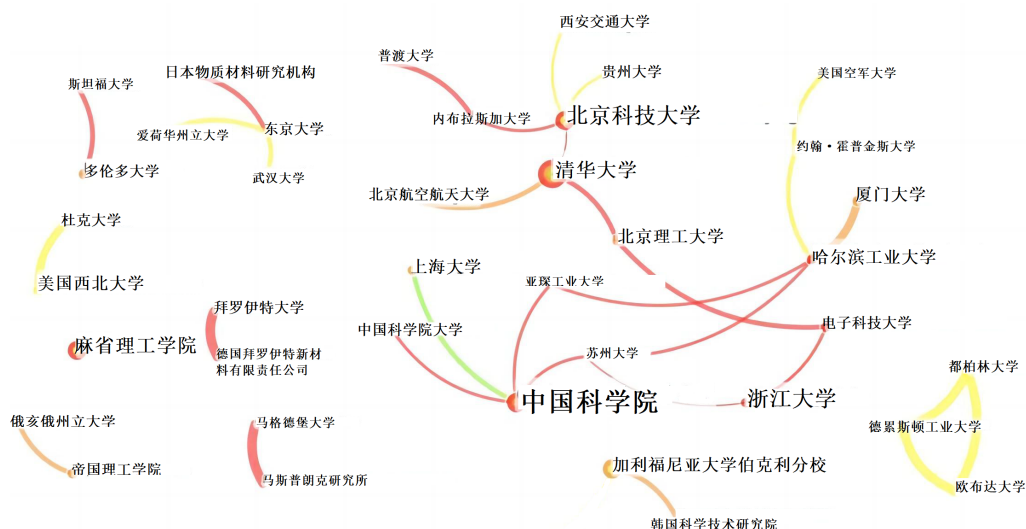


图4 研究机构合作关系

表 2 2012—2022 年 10 月国家发文量居于前 10 位分布情况

时间	第一作者	核心文献英文题目	刊名	本地数据源	网络总被引
				(纳入的 431 篇 高质量文献) 被引频次	频次 (次)
2018 年	Butler K T	Machine learning for molecular and materials science	《自然》(<i>Nature</i>)	53	1 372
2016 年	Raccuglia P	Machine-learning-assisted materials discovery using failed experiments	《自然》(<i>Nature</i>)	41	739
2013 年	Jain A	Commentary: the materials project: a materials genome approach to accelerating materials innovation	《应用物理快报 – 材料》 (<i>APL Materials</i>)	38	4 453
2018 年	Sanchez-Lengeling B	Inverse molecular design using machine learning: generative models for matter engineering	《科学》(<i>Science</i>)	32	639
2015 年	LeCun Y	Deep learning	《自然》(<i>Nature</i>)	30	21 320
2017 年	Liu Y	Materials discovery and design using machine learning	《无机材料科学学报》 (<i>Journal of Materiomics</i>)	28	500
2016 年	Ward L	A general-purpose machine learning framework for predicting properties of inorganic materials	《npj 计算材料学》 (<i>npj Computational Materials</i>)	24	628
2018 年	Gomez-Bombarelli R	Automatic chemical design using a data-driven continuous representation of molecules	《ACS 核心科学》 (<i>ACS Central Science</i>)	21	1 398
2018 年	Xie T	Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties	《物理评论快报》 (<i>Physical Review Letters</i>)	20	554
2016 年	Gomez-Bombarelli R	Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach	《自然 – 材料》 (<i>Nature Materials</i>)	20	512

通过总结分析关键核心节点文献,可以发现,这些文献集中涌现在 2013—2018 年,且均发表于高影响力期刊上,如《自然》《科学》等;均具有较高的被引频次,分别在理论发展和实证应用等方面奠定了坚实的基础,推动了人工智能在材料科学领域的应用发展。

其中,被引频次排名第一位的为 Butler K T 在 2018 年发表于《自然》的“Machine learning for molecular and materials science”,该综述报道了机器学习在化学科学中的应用进展,认为人工智能将加速分子和材料的设计、合成、表征和应用的发展,提出了前沿技术方向为:小数据集的知识挖掘、高效的化学数据处理表示方式、机器学习与量子计算结合、人工智能在材料科学领域的可解释性等^[8]。

被引频次排名第二位的为 Raccuglia P 于 2016 年发表的“Machine-learning-assisted materials discovery using failed experiments”,通过开发公共数据库收集大量未被发表的实验失败数据,明确了界定实验成功或失败所需的有效信息,将 3 955 条完整数据用于训练和测试支持向量机模型,并以决策树形式对其进行解释,最终以可验证假设的形式揭示了控制反应结果的化学原理,预测精度优于传统的人工分析,准确率达到 89%^[9]。被引频次排名第三位的为 Jain A 发表的“Commentary: the materials project: a materials genome approach to accelerating materials innovation”,重点介绍了作为材料基因组计划核心的材料项目,包括开放的数据集、数据挖掘、高通量计算、材料分析等,是具有成本效益和数据驱动

续表

序号	频次 (次)	关键词
13	15	遗传算法 (genetic algorithm)
14	15	人工神经网络 (artificial neural network)
15	12	微观结构 (microstructure)
16	8	高分子 (polymer)
17	7	吸附 (adsorption)
18	7	人工智能 (artificial intelligence)
19	6	模拟 (simulation)
20	6	卷积神经网络 (convolutional neural network)

2.6 关键词突现及时间线分析

通常情况下, 在某一研究领域内突然出现的一系列动态概念和潜在的研究问题, 往往能够反映该领域的研究前沿。因此, 在一定时间内, 以突现词为知识基础, 其所关联的多篇文献涵盖的研究主题, 可以作为判定和预测研究前沿的标准之一^[15]。本文使用 CiteSpace 软件开展突现词分析, 旨在研究人工智能在材料科学领域应用发展的研究前沿, 相关结果如图 6 所示, 共探究到 15 个突现词。

通过总结分析突现词时间范围与其代表的相关文献概要内容, 将这 15 个突现词划分为 3 个阶段, 总结归纳不同年度的人工智在材料领域相关的研究前沿。第一阶段为 2014—2017 年, 第二阶段为 2018—2020 年, 第三阶段为 2021—2022 年。

2014—2017 年包含的突现词主要为信息学、材料信息学、电导率、性能、复合材料和遗传算法, 该阶段是计算机科学、人工智能与材料科学研究领域融合发展理论研究的基础性阶段, 包括机器学习在内的数据科学迅速兴起, 构成了科学制造范式的第 4 个支柱, 并逐步衍生出在特种材料领域的研发应用研究。

2018—2020 年包含的突现词主要为蛋白质、氧化物、支持向量机、纳米颗粒、表层, 该阶段以机器学习为代表的人工智能技术在材料信息学领域得到了很好的应用发展, 研究对象包括药物分子、有机电子材料和晶体材料等, 涉及材料发现与设计、

性能预测与检测等研究, 研究方法包括高通量虚拟筛选、监督学习、支持向量机、反分子设计和贝叶斯优化等。

2021—2022 年包含的突现词为神经网络、深度神经网络、仿真、结构属性映射, 该阶段最突出的新兴研究前沿为深度学习算法在材料研发中的应用, 伴随着大数据技术、量子计算技术以及传感器技术的发展, 深度学习算法的应用得到了发展, 相较于传统机器学习算法, 其在预测精度、劳动力和计算成本以及人类经验依赖性等方面均展现出一定的优势。深度学习在图像识别、语音识别、自然语言处理和生物医学等领域表现出较好的应用效果; 在材料科学中, 卷积神经网络、递归神经网络、深度信念网络和深度编码网络等架构的深度学习算法在材料检测、材料分析、材料设计和量子化学等研究领域表现优异。

关键词时间线如图 7 所示。右侧的标号部分是关键词聚类结果, 共包含 8 类标签。分别为 #0 微观结构、#1 优化、#2 卷积神经网络、#3 设计、#4 金属-有机框架材料、#5 钙钛矿、#6 密度泛函理论和 #7 进化。左侧显示了每一类聚类集群内关键词基于时间轴的相互共现关系, 一定程度上展示了人工智能在材料科学领域应用发展的脉络与趋势。

对于每一类聚类集群, 选取频次排名前 5 位的关键词进行统计分析。其中, #0 微观结构集群包含的关键词为机器学习、预测、机械性能、微观结构、性能; #1 优化集群包含的关键词为优化、材料设计语言、遗传算法、复合材料、拓扑优化; #2 卷积神经网络集群包含的关键词为深度学习、人工智能、发现、分类、化学; #3 设计集群包含的关键词为设计、神经网络、人工神经网络、材料发现、数据库; #4 金属-有机框架材料集群包含的关键词为性能、仿真、贝叶斯优化、挑战、金属-有机框架材料; #5 钙钛矿集群包含的关键词为模型、聚合物、机械、强度、材料信息学; #6 密度泛函理论集群包含的关键词为密度泛函理论、机械弹性性能、晶体结构、氧化物、机械; #7 进化集群包含的关键词为进化、系统、一代、蛋白质、石墨烯细胞毒性。

总体而言, 人工智能在材料科学领域的应用发展脉络充分体现出了技术交叉融合推动科技创新发展的特性。随着研究人员在人工智能技术理论、共

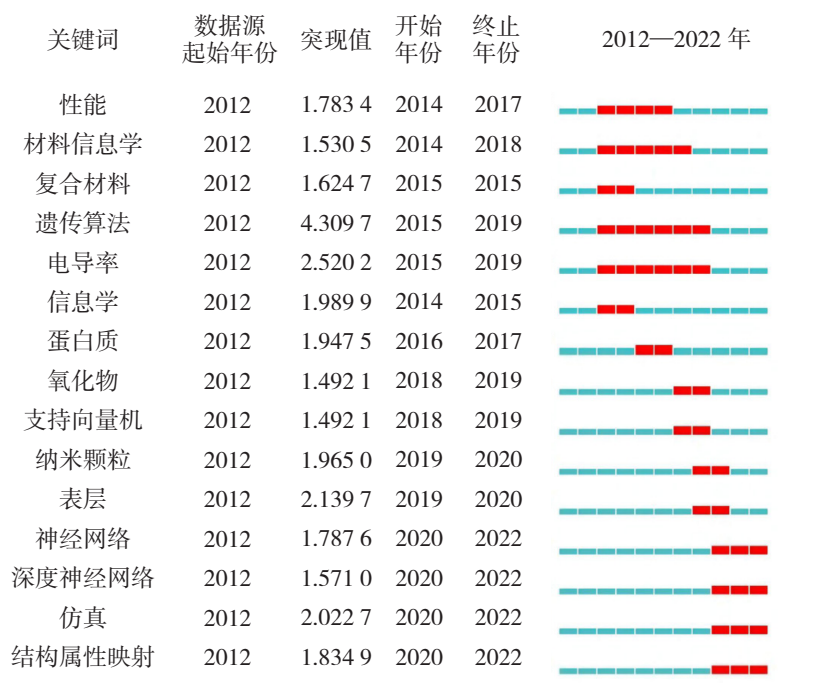


图 6 2012—2022 年 10 月关键词突现情况

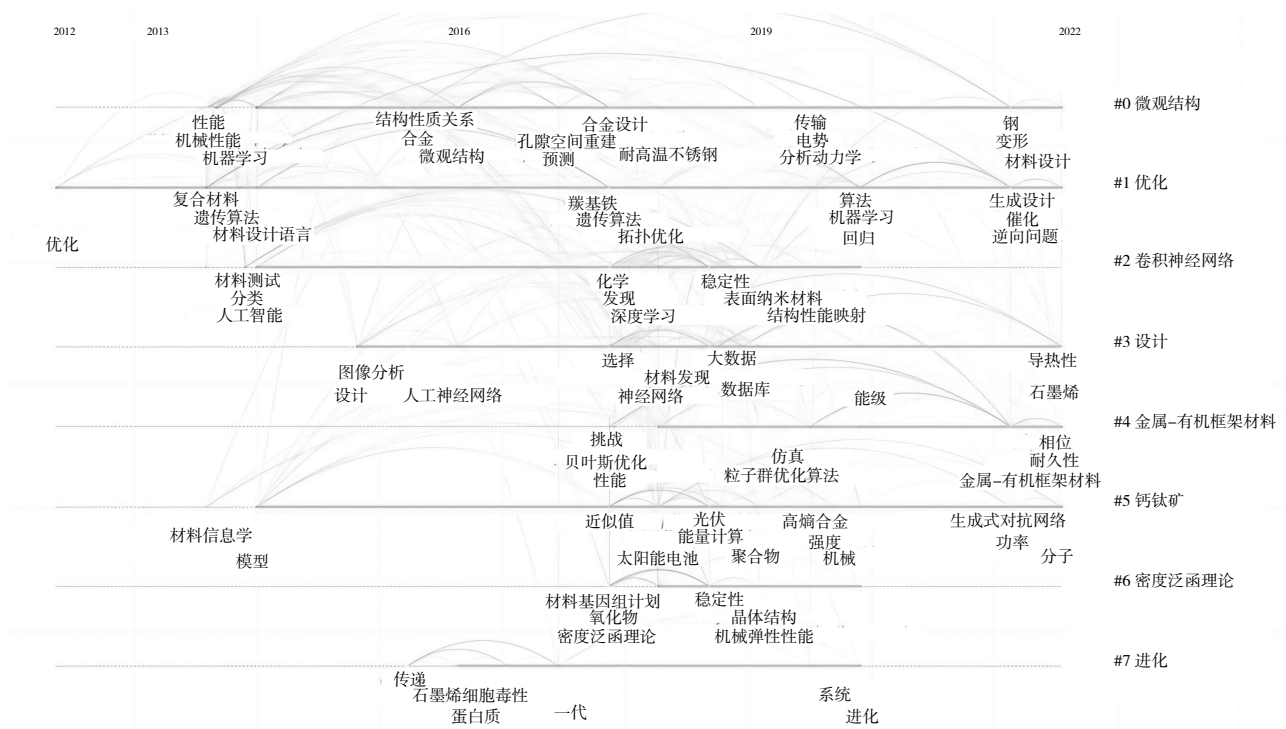


图 7 2012—2022 年 10 月关键词时间线

性关键技术等方面取得新突破,材料信息学的创新发展也得到了推动。以支持向量机、决策树和朴素贝叶斯等为代表的机器学习和以卷积神经网络、循环神经网络为代表的深度学习,逐渐成熟并被应用

于材料发现与设计、材料表征、材料检测与分析等多个领域。一方面,研究对象逐渐丰富,主要包括有机电子材料和晶体材料等;另一方面,研究方向逐步拓展,如材料发现与设计领域的结构导向设计、

元素设计、逆向设计和药物设计等, 材料检测分析领域的材料表征、材料降解分析、纳米材料分析和分子性能预测等。

3 结语

本文采用 CiteSpace 软件对 WOS 核心数据库收录的人工智能在材料科学领域应用相关的 431 篇高质量文献从时间分布、空间分布、研究机构分布、关键节点核心文献、关键词共现、关键词突现及时间线分析等方面进行分析, 得出以下结论。

3.1 总体情况

2012—2022 年人工智能在材料科学领域应用研究的发文量总体上呈现稳中有升的发展趋势, 且以 2018 年为分界点, 上升幅度变大。2012—2017 年是人工智能在材料科学研究领域的应用发展的基础性阶段。同时, 随着算力的进步, 人工智能技术实现突破性发展, 因此, 其在材料科学领域的研究成果随着时间的推移而持续快速增长, 说明某一关键技术的理论性突破或发展对科技创新具有强大的推动作用。2018 年及以后, 随着相关理论和共性技术的进步, 人工智能与材料科学领域融合发展逐步迈入应用阶段, 在药物、半导体材料研制等方面实现了模型创新和软件开发, 并逐步走向高计算效率、小数据集、可解释性和高预测精度等方面的优化研究。中国在该应用发展领域具备突出的研究活跃度和产出量, 发文量和中心性值均居首位, 且形成了节点数最多的研究机构合作集群。美国在国际合作方面具备较大优势, 跨国合作关系数量最多。

3.2 研究热点

本文综合分析关键节点核心文献、关键词共现、关键词突现及时间线图结果, 对人工智能在材料科学领域应用的研究热点进行归纳总结, 并借鉴典型人工智能研究流程进行研究热点的分类总结, 主要包括数据、模型以及应用 3 个方面。

数据方面, 高质量的数据可以避免信息的缺失或冗余, 从而更好地提升预测准确度, 其决定了人工智能在材料科学领域应用的“上限”。因此, 包括材料基因组计划等在内的材料项目均重点开展大型材料数据库的建设, 主要数据源为实验和模拟的材料物理性质、化学性质、结构性质、热力学性质以及动力学性质等材料性质数据, 反应速率和反

应温度等的化学反应数据, 材料的扫描电子显微镜图像、材料表面照片等的图像数据以及来自文献的数据。对于材料属性数据的表示方式亦是关键环节, 为人工智能算法模型提供输入, Fingerprints, SMILES (Simplified Molecular Input Line Entry System) 和加权图是常见的分子属性表示方法。

模型方面, 主要包括传统机器学习和深度学习两部分。支持向量机、决策树、随机森林、朴素贝叶斯、人工神经网络等多类型的机器学习算法均已得到了广泛应用。总体而言, 特定算法模型依据各自特点具有不同的应用范围。如支持向量机主要用于线性分类, 非常适合用于识别材料的结构属性映射关系, 已被用于对与靶药物相关的化合物进行分类; 朴素贝叶斯分类器的优点为只需要基于少量的训练数据来进行预测, 已用于预测新分子的合成配方。虽然传统机器学习算法已得到了广泛的应用, 但在材料数据特性表征与模型精度等方面, 深度学习算法均展示出一定的优势并逐步成为研究热点。各种架构的深度学习算法 (如卷积神经网络、递归神经网络等) 在材料检测、材料分析、材料设计和量子化学方面都表现出了优异的性能, 解决了由于材料数据的复杂性和标定难度导致的时间长和劳动力成本高等问题, 并提升了模型精度。

应用方面, 包括材料发现与设计、材料检测与分析以及量子化学等部分。材料发现主要为结构导向设计, 典型应用材料有钙钛矿等; 元素设计, 指从一组输入组合物中预测新的化合物及其结构; 逆向设计, 主要采用深度学习算法, 旨在实现以按需定制形式开展材料研发。材料检测与分析包括纳米材料分析、分子性质预测等。同时, 在量子化学领域, 人工智能算法可代替或结合计算机模拟开展复杂问题的计算, 能够节省研究成本、缩短研制周期。■

参考文献:

- [1] 谢曼, 干勇, 王慧. 面向 2035 的新材料强国战略研究 [J]. 中国工程科学, 2020, 22(5): 1-9.
- [2] 曾昆, 李晓芃, 沈紫云, 等. 我国新材料产业集群发展战略研究 [J]. 中国科学院院刊, 2022, 37(3): 343-351.
- [3] 王兴艳. 人工智能为材料工业发展赋能 [J]. 新材料产业, 2021(5): 49-52.
- [4] YAZDANI ASRAMI M, SADEGHI A, SONG W,

- et al. Artificial intelligence methods for applied superconductivity: material, design, manufacturing, testing, operation, and condition monitoring[J]. Superconductor science and technology, 2022, 35(12): 123001.
- [5] CHEN C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature[J]. Journal of the american society for information science and technology, 2005, 57(3): 359-377.
- [6] 王沛霖. 人工智能元年来临 [J]. 机器人产业, 2016(1): 2.
- [7] 郭乐倩, 李晶, 杨文方, 等. 基于 CiteSpace 进行人工智能、大数据与期刊融合发展的现状和热点分析 [J]. 新媒体研究, 2022, 8(2): 6-11.
- [8] BUTLER K T, DAVIES D W, CARTWRIGHT H, et al. Machine learning for molecular and materials science[J]. Nature, 2018, 559(7715): 547-555.
- [9] RACCUGLIA P, ELBERT K C, ADLER P D, et al. Machine learning-assisted materials discovery using failed experiments[J]. Nature, 2016, 533(7601): 73-76.
- [10] ANUBHAV J, SHYUE P O, GEOFFROY H, et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation[J]. APL materials, 2013(1): 011002.
- [11] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7753): 436-444.
- [12] LIU Y, ZHAO T, JU W, et al. Materials discovery and design using machine learning[J]. Journal of materiomics, 2017, 3(3): 159-177.
- [13] WARD L, AGRAWAL A, CHOUDHARY A, et al. A general-purpose machine learning framework for predicting properties of inorganic materials[J]. Npj computational materid, 2016, 2(1): 70-76.
- [14] 侯剑华, 胡志刚. CiteSpace 软件应用研究的回顾与展望 [J]. 现代情报, 2013, 33(4): 99-103.
- [15] KLEINBERG J. Bursty and hierarchical structure in streams[J]. Data mining and knowledge discovery, 2003, 7(4): 373-397.

Research of the Application and Development of Artificial Intelligence in Material Science Field: A Bibliometric Analysis Based on WOS Core Database

CHU Xulong^{1,2}, LIU Jinchang¹, SHI Dongmei¹

(1. High Technology Research and Development Center, Ministry of Science and Technology of the People's Republic of China, Beijing 100044;

2. School of Public Administration, Beihang University, Beijing 100191)

Abstract: The traditional methods for discovering new materials, including experimental trial-and-error methods and methods based on density functional theory, gradually show shortcomings in development cost, efficiency and cycle. Artificial intelligence, especially machine learning and deep learning, combining powerful data processing and high predictive performance, demonstrates advantages in material discovery, material design, material detection, and material analysis, and plays an important role in promoting innovation in new material research and development. This paper uses relevant articles collected from the Web of Science (WOS) core database under the Institute of Information and Intelligence (ISI) in the past 10 years as the data source, and uses CiteSpace software for visual analysis, aiming to summarize and analyze the research status, hotspots, and the trends of artificial intelligence in the field of materials science, and to provide relevant reference for the integrated innovation and development of artificial intelligence and materials science in China, and improve its research level.

Keywords: artificial intelligence; interdisciplinary science; material science; bibliometrics