doi:10.3772/j.issn.2095-915x.2026.04.008

基于 BERTopic 和 ARIMA 的文献热点主题及演化趋势分析
—— 以 ESI 材料科学为例

开放科学（资源服务）标识码（OSID）

许欣

江苏科技大学苏州理工学院图书馆 苏州 215600

摘要：[目的 / 意义] 本研究基于深度学习的主题建模方法，分析 ESI 材料科学领域的热点主题并预测其演化趋势，为科研人员把握学科动态，为决策者优化资源配置提供参考。[方法 / 过程] 利用 BERTopic 模型对 2014—2023 年 ESI 材料科学领域的 12138 篇高被引文献进行主题识别，并用 ARIMA 模型对各主题的文献数量进行时间序列预测分析。[结果 / 结论] 研究结果表明，近十年 ESI 材料科学领域的热门主题主要集中在电化学储能、可再生能源的催化反应、柔性可穿戴智能传感器系统、生物材料的医学应用、太阳能电池等方面，充分体现出理工融合、医工融合的特点。在当下及未来一段时间内，ESI 材料科学领域的新能源转化与存储、医学器材等研究主题都将受到广泛关注。

关键词：BERTopic 模型；主题识别；ARIMA；时间序列模型

中图分类号：G350.7；G353；TP393

Analysis of Hot Topics and Evolution Trends of Literature Based on BERTopic and ARIMA: A Case Study of ESI Materials Science Field

XU Xin

Library of Suzhou Institute of Technology, Jiangsu University of Science and Technology, Suzhou 215600, China

Abstract: [Purpose/Significance] This paper analyzes the hot topics in the ESI Materials Science field and predicts their evolution trends based on deep learning topic modeling methods, providing references for researchers to grasp disciplinary dynamics and decision-makers to optimize resource allocation. [Method/Process] The BERTopic model was used to identify topics from 12138 highly cited papers in the ESI Materials Science field from 2014 to 2023, and the ARIMA model was used for time-series forecasting analysis of the number of papers for each topic. [Result/Conclusion] The research results show that the hot topics in the ESI Materials Science field over the past decade mainly focus on electrochemical energy

作者简介 许欣（1996-），通信作者，硕士，助理馆员，主要研究方向为学科服务，E-mail：xuxin2019@163.com。

引用格式 许欣 . 基于 BERTopic 和 ARIMA 的文献热点主题及演化趋势分析 —— 以 ESI 材料科学为例 [J]. 情报工程 , 2026, 12(4): 93-107.

doi:10.3772/j.issn.2095-915x.2026.04.008

storage, catalytic reactions for renewable energy, flexible wearable intelligent sensor systems, medical applications of biomaterials, solar cells, etc., fully reflecting the characteristics of Interdisciplinary subject. In the present and near future, research in the ESI Materials Science field related to new energy conversion and storage and medical devices will continue to receive extensive attention.

Keywords: BERTopic Model; Topic Identification; ARIMA; Time Series Model

引言

进入 21 世纪，材料科学领域的前沿研究方向逐渐朝着能量转换与存储材料、生物医用与仿生材料、纳米材料与器件、环境友好材料等方向发展 [1]。随着材料科学在能源、医疗、电子等领域的交叉融合不断深化，相关研究对跨学科协同的需求日益凸显。学科研究主题的多元化发展，也使得高校图书馆积极开展学科研究主题分析工作，成为近年来提升学科竞争力的重要研究方向 [2]。准确识别材料科学领域的热门研究主题，有助于科研工作者全面把握领域研究现状、特征及发展趋势，为科研实践提供参考，同时也能为高校材料学科的建设规划提供理论指导 [3]。

近年来，国内外对学科主题的识别和分析已有丰富的研究成果。黄东 [4] 以图书情报领域科研文献为研究对象，构建高频关键词共词网络，梳理了 2016—2020 年的研究主题。Wu 等 [5] 基于 LDA 模型识别不同时间的主题，来分析石墨烯领域专利文献的主题演化路径；李俊炀等 [6] 基于 LDA 模型对图情领域近 30 年关于虚拟现实技术的文献摘要进行主题分析，研究了各个主题的热点及趋势。基于共词网络分析的方法和以 LDA 为代表的传统主题模型在分析过程中往往将高频词汇的共现规律作为依据，对词语所处的上下文语境及其具体含义理解不够，在解读文档文本的深层语义和上下文逻辑关联时的解释不足。BERTopic 模型能够生成可解释性高的主题，近年来备受关注。BERTopic 是由 Grootendorst[7] 提出的一种融合 BERT 预训练模型与主题建模的语义分析工具，用 BERT 模型提取文本语义向量，并通过主题建模算法对语义向量进行聚类分析，实现对文本的语义分析。已有学者将该模型应用于多种文本主题分析，如政府工作报告 [8]、教育政策 [9]、科技人才政策 [10] 等领域的研究。在学科主题分析方面，张家惠等 [11] 以 2013—2022 年数据安全领域论文的标题和摘要为文本数据，借助 Python 的 BERTopic 库开展主题识别与解读；杨思洛等 [12] 运用 BERTopic 模型，对 2019—2023 年国内信息资源管理期刊论文进行了主题识别；徐汉青 [13] 则收集了 1995—2024 年 LIS 学科中 AI 研究文献的标题、关键词及摘要，构建语料库并分析该领域研究主题的演变趋势。这些研究表明，BERTopic 模型在学科主题分析中具有较高的可行性。在主题的演化方面，已有研究多为以人工判读的方式来分析学科主题时序变化的趋势，通过构建科学严谨的数理模型来进行时序预测的研究较少。因此，本文尝试运用 BERTopic 模型结合自回归移动平均模型（Autoregressive Integrated Moving Average Model,

ARIMA）的方法，对全球材料科学领域的研究主题及其演化趋势进行深入分析，以期为材料科学的科研工作和学科建设提供有益参考。

1 研究思路与方法

1.1 研究思路

本文基于 BERTopic 和 ARIMA 识别文献的主题并定量研究各主题的演变过程及发展趋势。具体研究步骤如下：（1）数据收集与预处理。通过 Incites 平台筛选学术文献并从 Web of Science 核心合集数据库中收集文献的标题、摘要、发表日期等信息。研究选择 ESI 材料科学 Q1 期刊的高被引论文作为研究对象，聚焦高水平、高影响力研究成果，把握研究热点。（2）主题模型构建。使用 BERTopic 模型确定最优主题数量，并通过层次聚类识别研究方向，以揭示 ESI 材料科学研究的主要热门主题。（3）主题演化趋势分析。基于文献发表时间，计算不同年份各主题文献数量，并基于 ARIMA 模型预测主题的未来发展趋势，以可视化的形式呈现。

1.2 基于BERTopic模型的主题识别

利用 BERTopic 模型对材料科学领域文献进行主题识别。BERTopic 是基于 BERT 词向量的主题建模技术，通过 Transformers 架构生成深层语义表征，实现对主题的解释分析。该模型的主题建模分析流程可以分为嵌入（Embeddings）、降维（Dimensionality Reduction）、聚类（clustering）、序列化（Tokenizer）、加权（Weighting Scheme）五大核心步骤。

首先提取论文摘要，整理成标准 .txt 文件格式。使用 BERTopic 默认的 Sentence-Transformers 方法进行文本嵌入。在降维环节，使用 UMAP（Uniform Manifold Approximation and Projection for Dimension Reduction）降维方法，继而通过 HDBSCAN（Hierarchical Density-Based Spatial Clustering of Applications with Noise）算法进行聚类。在序列化步骤选择 CountVectorizer 特征数值化方法，再利用 c-TF-IDF 算法进行加权，最后进行主题识别。

1.3 基于ARIMA的时间序列预测

自回归移动平均模型，又称 Box-jenkins 模型，由 Box 和 Jenkins 在 20 世纪 70 年代初提出，是统计学与计量经济学等领域广泛应用的经典时间序列分析方法，也是时间序列预测中最常见的模型之一。ARIMA 模型主要由三部分构成，分别为自回归模型（AR）、差分过程（I）和移动平均模型（MA）。其中自回归模型（AR）处理时间序列的自回归部分，它考虑过去的观测值对当前值的影响；差分过程（I）通过一阶或者二阶等差分处理，使非平稳时间序列达到平稳；移动平均模型（MA）处理时间序列的移动平均部分，它考虑过去的预测误差对当前值的影响。

ARIMA 模型由 3 个重要参数 p、d、q 决定。p 为自回归系数（Auto-Regressive），表示观测值滞后 p 阶；d 对应 ARIMA 模型中的 “I”，是差分的阶数，表示时间序列变为平稳需要至少进行 d 次差分；q 即滑动平均系数（Moving Average），表示误差项滞后 q 阶。对时间序列做自相关图平稳检验，分别查看 ACF 图和 PACF

doi:10.3772/j.issn.2095-915x.2026.04.008

图，并对非平稳序列再进行差分，直到满足平稳性条件，则 d 值确定。序列平稳之后，参数 p、q 的确定方式如表所列。

表 1　ARIMA 模型参数 p、q 的确定方式

模型	ACF	PACF
AR（p）模型	拖尾	p 阶截尾
MA（q）模型	q 阶截尾	拖尾
ARMA（p, q）模型	q 阶拖尾	p 阶拖尾

2　数据收集与预处理

通过 Incites 平台筛选数据，时间分别选择 2014—2023 年，学科分类体系选择“Essential Science Indicators”，研究领域是“Materials Science”，期刊影响因子（JIF）分区选择“Q1”，文献类型限定为“Article，Review”。将筛选结果“在 Web of Science 中查看”，新增限定“高被引论文”，得到文献 12252 篇。剔除缺少摘要信息的文献，最终选取 12138 篇材料科学领域 Q1 的高被引文献进行后续分析。

3　ESI 材料科学文献热点主题识别分析

使用 BERTopic 模型进行主题建模时，在综合对比多次建模结果的基础上，最终确定 UMAP 的参数设置为：n_neighbors（临近样本点数量）=15，n_components（降维空间维数）=2，min_dist=0.0，metric=’cosine’，random_state=42。每个聚类最小的文章数量设置为 50，ngram 范围为（1,3）。在设定了上述参数的条件下，模型共识别出 42 个材料科学领域的研究主题（Topic0 至 Topic41），涵盖数据集中的 9460 篇文献摘要，另有 2678 篇文献摘要被分为噪音数据（Topic-1）。噪音数据包含少数离群点论文，还有聚类规模没有达到设定参数（聚类的文章数量小于 50）的研究主题。在此基础上，结合主题层次聚类法对获取的 42 个研究主题进行分组，获取材料科学领域的 9 组研究方向。

	Topic	Count	Name	Representation	Representative_Docs
0	-1	2678	-1_materials_properties_mechanical_devices	[materials, properties, mechanical, devices, t...	[The sensing of bioactive molecules based on p...
1	0	1395	0_batteries_lithium_electrolyte_battery	[batteries, lithium, electrolyte, battery, ion...	[Lithium-sulfur (Li-S) batteries are considere...
2	1	1024	1_reaction_catalysts_hydrogen_electrocatalysts	[reaction, catalysts, hydrogen, electrocatalys...	[Developing active and stable electrocatalysts...
3	2	696	2_cancer_nanoparticles_cells_photothermal	[cancer, nanoparticles, cells, photothermal, i...	[Insufficient oxygenation (hypoxia), acidic pH...
4	3	535	3_sensors_wearable_sensor_skin	[sensors, wearable, sensor, skin, stretchable,...	[High-performance flexible strain and pressure...
5	4	473	4_photocatalytic_photocatalysts_c3n4_photocata...	[photocatalytic, photocatalysts, c3n4, photoca...	[With the rapid development of economy and mod...
6	5	453	5_batteries_electrolyte_zinc ion_zn2	[batteries, electrolyte, zinc ion, zn2, ion ba...	[Zn dendrites growth and poor cycling stabilit...
7	6	267	6_manufacturing_printing_additive manufacturin...	[manufacturing, printing, additive manufacturi...	[Additive manufacturing, often known as three-...
8	7	264	7_alloys_entropy_high entropy_alloy	[alloys, entropy, high entropy, alloy, entropy...	[High-entropy alloys (HEAs) have displayed num...
9	8	248	8_perovskites_halide perovskites_optoelectroni...	[perovskites, halide perovskites, optoelectron...	[HighlightsLead-based halide perovskite materi...
10	9	246	9_hydrogels_hydrogel_strain_self healing	[hydrogels, hydrogel, strain, self healing, co...	[Flexible and wearable hydrogel strain sensors...
11	10	209	10_cellulose_self healing_materials_coatings	[cellulose, self healing, materials, coatings,...	[Organic coatings are one of the most popular ...

图 1　BERTopic 模型生成的主题聚类结果

根据各聚类的代表性关键词及对应文献的摘要文本，总结出 9 个分组的研究主题，根据文献数量排序分别是 Group1（电化学储能）、Group2（可再生能源的催化反应）、Group3（柔性可穿戴智能传感器系统）、Group4（多功能复合材料）、Group5（生

物材料的医学应用）、Group6（材料性能研究）、Group7（光电器件中材料的选择和应用）、Group8（太阳能电池）、Group9（二维材料）。

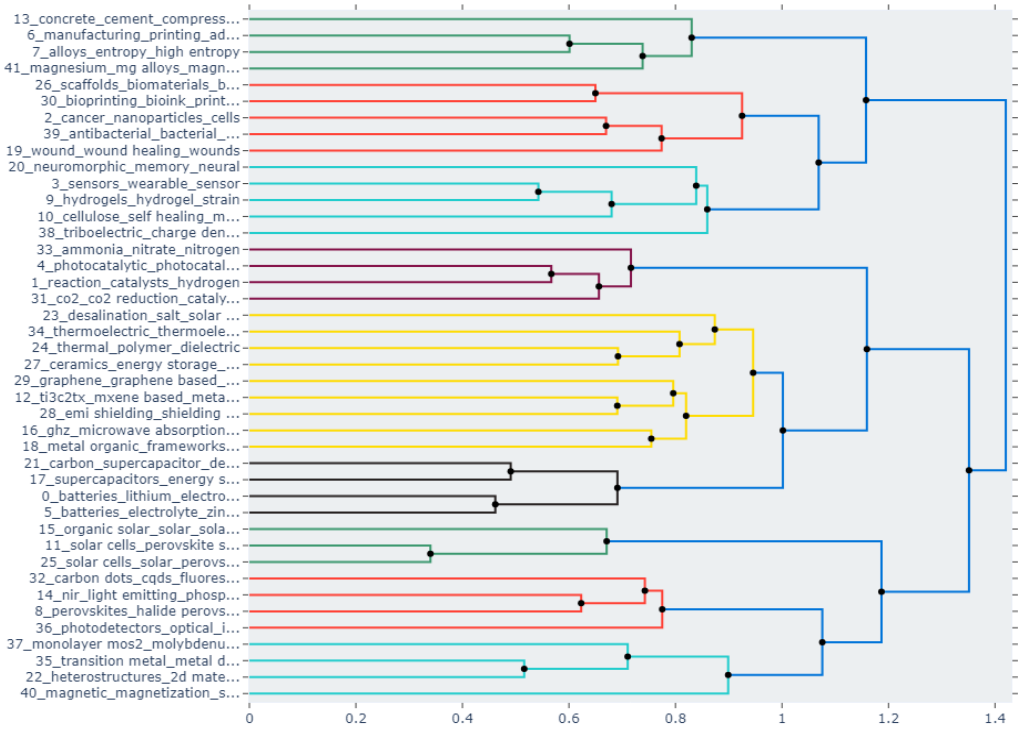


图 2 BERTopic 模型主题层次聚类结果

表 2 ESI 材料科学热点主题

主题组	主题 聚类	代表性关键词	文献数量
Group1	Topic0	batteries, lithium, electrolyte, battery, ion batteries, electrolytes, storage, lithium ion, electrochemical, energy density	1395
	Topic5	batteries, electrolyte, zinc ion, zn^{2+}, ion batteries, zinc ion batteries, zn anode, aqueous zinc, battery, energy storage	453
	Topic17	supercapacitors, energy storage, supercapacitor, electrochemical, electrode materials, capacitance, energy storage devices, power density, batteries, high energy	169
	Topic21	carbon, supercapacitor, density, energy, $nico_2s_4$, electrochemical, storage, specific capacitance, surface area, biomass	147
Group2	Topic1	reaction, catalysts, hydrogen, electrocatalysts, evolution reaction, catalyst, catalytic, water splitting, oxygen evolution, oxygen evolution reaction	1024
	Topic4	photocatalytic, photocatalysts, c_3n_4, photocatalyst, photocatalysis, solar, hydrogen, co_2 reduction, visible light, water splitting	473
	Topic31	co_2, co_2 reduction, catalysts, electrochemical co_2, electrochemical co_2 reduction, electrocatalytic, reduction reaction, reduction reaction co_2rr, reaction co_2rr, faradaic efficiency	85
	Topic33	ammonia, nitrate, nitrogen, faradaic efficiency, catalysts, electrochemical, reduction reaction nrr, reaction nrr, bosch process, nh_3 yield	77

doi:10.3772/j.issn.2095-915x.2026.04.008

（续表 2）

主题组	主题 聚类	代表性关键词	文献数量
Group3	Topic3	sensors, wearable, sensor, skin, stretchable, pressure sensors, tactile, soft, high sensitivity, textile	535
	Topic9	hydrogels, hydrogel, strain, self healing, conductive hydrogels, wearable, tissue, mechanical properties, gels, polymer	246
	Topic10	cellulose, self healing, materials, coatings, wood, surface, composites, polymers, nanocellulose, fiber	209
	Topic20	neuromorphic, memory, neural, synapses, memristors, switching, machine learning, neuromorphic computing, plasticity, neural networks	149
	Topic38	triboelectric, charge density, nanogenerator, triboelectric nanogenerator, triboelectric nanogenerators, nanogenerators, contact electrification, triboelectric nanogenerator teng, nanogenerator teng, energy harvesting	60
Group4	Topic12	$ti_3c_2t_x$, mxene based, metal carbides, transition metal, ti_3c_2, transition metal carbides, energy storage, materials, metal, 2d transition metal	200
	Topic16	ghz, microwave absorption, dielectric, reflection loss, absorbers, carbon, absorption bandwidth, electromagnetic wave, emw absorption, impedance	169
	Topic18	metal organic, frameworks, organic frameworks, metal organic frameworks, frameworks mofs, organic frameworks mofs, mof based, porous, structures, mof derived	160
	Topic23	desalination, salt, solar steam, evaporator, solar steam generation, evaporation rate, solar energy, evaporators, water evaporation, solar thermal	142
	Topic24	thermal, polymer, dielectric, composites, thermal conductivity, polymer composites, thermally conductive, nanocomposites, thermal management, boron nitride	127
	Topic27	ceramics, energy storage, capacitors, lead free, energy storage density, storage density, energy storage performance, storage performance, antiferroelectric, recoverable energy	99
	Topic28	emi shielding, shielding materials, emi shielding materials, interference emi, electromagnetic interference emi, interference emi shielding, shielding effectiveness, $ti_3c_2t_x$, conductive, polymer	98
	Topic29	graphene, graphene based, materials, graphene oxide, oxide, thermal, exfoliation, electrical, reduced graphene, nanocomposites	94
	Topic34	thermoelectric, thermoelectric materials, thermal, thermal conductivity, thermoelectrics, lattice thermal, lattice thermal conductivity, te materials, thermoelectric performance, bi_2te_3	74
Group5	Topic2	cancer, nanoparticles, cells, photothermal, immunotherapy, nanoparticle, tumour, nanomaterials, drug delivery, cancer therapy	696
	Topic19	wound, wound healing, wounds, hydrogels, adhesive, wound dressing, wound closure, wound dressings, diabetic wound, self healing	151
	Topic26	scaffolds, biomaterials, bone tissue, tissue engineering, implant, bone regeneration, bone tissue engineering, bone defects, macrophages, osteogenic	105
	Topic30	bioprinting, bioink, printing, tissue engineering, bioinks, biomaterials, regenerative, 3d bioprinting, stem cell, biofabrication	87
	Topic39	antibacterial, bacterial, antimicrobial, antibiotic, antibiotics, infection, biofilms, resistant, antibacterial agents, wound	60
Group6	Topic6	manufacturing, printing, additive manufacturing, 3d printing, microstructure, fatigue, materials, alloy, laser melting, mechanical properties	267
	Topic7	alloys, entropy, high entropy, alloy, entropy alloys, high entropy alloys, entropy alloy, strength ductility, mechanical properties, high entropy alloy	264
	Topic13	concrete, cement, compressive strength, recycled, durability, fly ash, carbonation, geopolymer, hydration, elsevier ltd rights	188
	Topic41	magnesium, mg alloys, magnesium alloys, alloy, corrosion resistance, high strength, magnesium mg, mg alloy, alloying, magnesium alloy	53

（续表 2）

主题组	主题 聚类	代表性关键词	文献数量
Group7	Topic8	perovskites, halide perovskites, optoelectronic, nanocrystals, metal halide, light emitting diodes, emitting diodes, diodes, photoluminescence, lead halide	248
	Topic14	nir, light emitting, phosphors, emitters, luminescence, fluorescence, light emitting diodes, emitting diodes, organic light emitting, organic light	182
	Topic32	carbon dots, cqds, fluorescence, emission, quantum dots, photoluminescence, dots cds, carbon dots cds, bioimaging, color	78
	Topic36	photodetectors, optical, infrared, pds, photodetector, photodetection, optoelectronic, detectivity, dark current, polarization	71
Group8	Topic11	solar cells, perovskite solar, cells, perovskite solar cells, perovskites, power conversion efficiency, devices, photovoltaic, perovskite films, charge	200
	Topic15	organic solar, solar, solar cells, organic solar cells, polymer, photovoltaic, non fullerene, power conversion efficiency, solar cells oscs, cells oscs	173
	Topic25	solar cells, solar, perovskite solar, cells pscs, solar cells pscs, perovskite solar cells, sno_2, power conversion efficiency, conversion efficiency pce, devices	117
Group9	Topic22	heterostructures, 2d materials, transistors, van der waals, silicon, semiconductors, quantum, graphene, semiconductor, wse2	143
	Topic35	transition metal, metal dichalcogenides, mos_2, transition metal dichalcogenides, metal, two dimensional, ws_2, metal dichalcogenides tmds, dichalcogenides tmds, sapphire	72
	Topic37	monolayer mos_2, molybdenum disulfide, layer mos_2, molybdenum disulfide mos_2, disulfide mos_2, exfoliation, mos_2 nanosheets, $mose_2$, two dimensional, 2h phase	61
	Topic40	magnetic, magnetization, spin orbit, ferromagnetic, spintronics, magnetic field, magnetism, orbit torque, spin orbit torque, antiferromagnets	59

（1）Group1 电化学储能

Group1 包括 Topic0、Topic5、Topic17 和 Topic21。Topic0 的代表性关键词有 batteries、lithium、electrolyte 等，主要围绕锂离子电池展开。Topic5 是锌离子电池相关的主题，主要关键词有 electrolyte、zinc ion batteries 等。Topic17 是关于超级电容器的主题，代表性关键词包括 supercapacitors、energy storage 等。Topic21 则是与碳材料有关的主题，代表性关键词有 carbon、supercapacitor、density 等。

这一组的主题主要涉及储能技术，包括电池和超级电容器。锂离子和锌离子电池是电化学储能的重要研究方向；超级电容器作为高功率设备，常与电池结合使用；而碳材料则可作为电池和超级电容器的关键电极材料 [14]。

（2）Group2 可再生能源的催化反应

Group2 包括 Topic1、Topic4、Topic31 和 Topic33，主要研究的是可再生能源的催化反应。Topic1 第一组属于能源转换领域，主要是关于电催化水分解的，有 reaction、catalysts、hydrogen 等代表性关键词。Topic4 的代表性关键词有 photocatalytic、photocatalysts、C_3N_4 等，属于光催化和太阳能利用的研究主题。Topic31 是二氧化碳电催化还原的研究，代表性关键词有 Co_2 reduction、catalysts 等。Topic33 是电催化硝酸盐还原合成氨的主题，代表性关键词有 ammonia、nitrogen、faradaic efficiency 等。

（3）Group3 柔性可穿戴智能传感器系统

Group3 包括 Topic3、Topic9、Topic10、Topic20 和 Topic38。Topic3 的主要关键词有

doi:10.3772/j.issn.2095-915x.2026.04.008

pressure sensors、skin、stretchable 等，关注的是与皮肤接触的柔性传感器。Topic9 的主要关键词有 hydrogels、strain、self healing、wearable 等。水凝胶作为生物相容性材料，具有自修复和导电特性，可用于生物医学领域。Topic10 包括 cellulose、self healing、coatings、wood、surface 等主要关键词。纤维素常用于制备自修复涂层和复合材料，可应用于柔性电子或生物医学领域。Topic20 包括 neuromorphic、synapses、memristors、switching、machine learning 等主要关键词，涉及神经形态计算，模仿生物神经网络的硬件，用于人工智能和机器学习。Topic38 的主要关键词有 triboelectric、charge density、nanogenerator 等。摩擦电纳米发电机是一种将机械能转化为电能的技术，常用于自供电传感器或能量收集设备。

导电性水凝胶在电子皮肤和个性化医疗监测方面具有巨大潜力 [15]，而纤维素处理得到的纤维素衍生物可以均匀地分散在水中或溶剂中 [16]，这使得它们对开发生物医学材料具有吸引力，包括应用于柔性可穿戴传感器。神经形态计算可以用于处理传感器数据，提高设备的智能性；摩擦电纳米发电机则可为这些可穿戴设备提供可持续和普遍的能源解决方案 [17]。这一组主题的文献主要围绕开发柔性可穿戴智能传感器所涉及的各个组成部分，体现了材料科学与生物医学、人工智能等其他科学领域的交叉融合。

（4）Group4 多功能复合材料

Group4 包括 Topic12、Topic16、Topic18、Topic23、Topic24、Topic27、Topic28、Topic29 和 Topic34。Topic12 主要涉及 MXene 材料，特别是 $Ti_3C_2T_x$，属于二维过渡金属碳化物，常用于储能和导电材料。Topic16 主要关键词包括 ghz、microwave absorption、dielectric、reflection loss 等，属于电磁波吸收领域。Topic18 主要是金属有机框架（MOFs），多用于吸附或催化。Topic23 有 desalination、salt、solar steam、evaporator 等主要关键词，涉及太阳能利用和水蒸发。Topic24 是聚合物复合材料及其导热性。Topic27 主要是围绕 ceramics 和 energy storage 的研究。Topic28 是 EMI 屏蔽材料、导电聚合物、Ti3C2Tx 的应用。Topic29 围绕 graphene 展开，涉及 thermal、exfoliation、electrical 等。Topic34 是热电材料的研究，包括 thermal conductivity、thermoelectric performance 等。

综合来看，这组主题主要围绕多功能复合材料，用于能源存储、电磁波管理、热管理、太阳能利用等多个领域。例如含有 BNNSs 的先进聚合物纳米复合材料 [18]，同时具有高导热性、优异的电绝缘性和出色的柔韧性，应用于现代电气系统和电子设备的热管理；或提升了容量、倍率能力和循环稳定性的 3D MXene 薄膜 [19]，可以指导未来为储能、催化、环境和生物医学的应用设计高性能 MXene 基材料。该组主题体现了材料科学、能源工程与环境科学的深度交叉，强调通过材料创新解决实际问题。

（5）Group5 生物材料的医学应用

Group5 包括 Topic2、Topic19、Topic26、Topic30 和 Topic39，主要涉及材料科学在医学领域的应用。Topic2 是关于癌症治疗的主题，包括 cancer、nanoparticles、cells 等。Topic19 是关于伤口愈合的主题，主要关键词有 wound healing、hydrogels 等。Topic26 是支架和骨组织

工程，代表性关键词有 scaffolds、biomaterials、bone regeneration 等。Topic30 是生物打印的主题，包括 bioprinting、bioink、tissue engineering 等。Topic39 是抗菌材料的主题，比如 antibacterial、antimicrobial, antibiotic、antibiotics 等。

（6）Group6 材料性能研究

Group6 包括 Topic6、Topic7、Topic13 和 Topic41。Topic6 的研究方向是增材制造工艺对材料性能的影响，主要关键词有 additive manufacturing、3d printing、microstructure 等。Topic7 研究的是高熵合金的机械性能，关键词包括 high entropy alloy、strength ductility、mechanical properties 等。Topic13 围绕混凝土展开，主要关键词有 concrete、cement、compressive strength 等。Topic41 的主要关键词有 magnesium alloys、corrosion resistance、high strength 等，集中在镁合金的研究。

这组主题主要针对各类材料的化学性能和机械性能展开研究，比如合金材料的强度、耐疲劳性、耐腐蚀性，混凝土材料的耐久性等。

（7）Group7 光电器件中材料的选择和应用

Group7 包括 Topic8、Topic14、Topic32 和 Topic36。Topic8 的研究围绕钙钛矿展开，核心关键词有 perovskites、halide perovskites、optoelectronic 等。Topic14 是关于发光材料的研究，核心关键词包括 nir、light emitting、phosphors、luminescence 等。Topic32 关注碳基纳米材料，如碳点和量子点在荧光发射、生物成像方面的应用，有 carbon dots、cqds、fluorescence 等主要关键词。Topic36 的研究围绕光电探测器展开，代表性关键词有 photodetectors、optical、infrared 等。

这组主题研究的是光电器件中材料的选择和应用。发光材料和光电探测器都是光电器件相关领域的研究对象，卤化物钙钛矿和碳点因其优异的光学和电子性能也被广泛研究应用于光电器件 [20-21]。

（8）Group8 太阳能电池

Group8 包括 Topic11、Topic15 和 Topic25，共同的主题是太阳能电池。Topic11 关注钙钛矿太阳能电池的性能和结构，关键词有 perovskite solar cells、power conversion efficiency、devices 等。Topic15 涉及有机太阳能电池，代表性关键词有 polymer、photovoltaic、non fullerene 等。Topic25 围绕钙钛矿太阳能电池的效率和稳定性展开研究，主要关键词有 solar cells、perovskite solar、cells pscs 等。

（9）Group9 二维材料

Group9 包括 Topic22、Topic35、Topic37 和 Topic40。Topic22 围绕 van der waals 展开研究，关键词有 heterostructures、2dmaterials、transistors 等。Topic35 主要是涉及二维过渡金属二硫化物的研究，关键词有 MoS_2、transition metal dichalcogenides 等。Topic37 进一步细化了金属二硫化物的不同形态和制备方法，主要关键词有 layer MoS_2、molybdenum disulfide MoS_2、exfoliation 等。Topic40 涉及材料的磁性特性，代表性关键词有 magnetic、magnetization、spin orbit 等。

该主题主要是面向不同类型二维材料的研究，包括性质、制备技术、应用等方面，还有自旋现象和磁性在二维材料中的研究。例如有基于二维材料的范德华异质结构中超快电荷转

doi:10.3772/j.issn.2095-915x.2026.04.008

移可以实现用于光电子和光收集的新型二维器件的研究；或是利用 MoS_2 的光致发光效应分析其应变与缺陷的研究等。

4 ESI 材料科学文献热点主题演化趋势分析

为了更好把握材料科学研究主题的演化趋势，对 BERTopic 主题模型的训练结果中获取的 9 组主题（42 个 Topic），使用 ARIMA 时间序列模型绘制 2014—2023 十年间的时间序列图，并预测 2024—2026 年每个 Topic 文献数量的趋势。

以 Group1 的 Topic0 为例。将数据导入 IBM SPSS Statistics（version R27.0.1.0）软件，通过原始序列、AC 和 PACF 检验，发现原始序列图并未见明显平稳特征，故尝试采用一阶差分处理。处理后 ACF 和 PACF 检测分别为 1 阶拖尾和 0 阶拖尾，故一阶差分后的序列是平稳序列，p、d、q 分别是 1、1、0。

再使用 SPSS 软件中的“专家建模器”进行建模，得到的模型是 ARIMA(0,1,0)。对 ARIMA(1,0,0)、ARIMA(1,1,0) 和 ARIMA(0,1,0) 分别建模并对模型统计的 R 方值和残差检验进行比较，最终选择 ARIMA(1,1,0) 模型。R 方达到 0.779，如表 3 所示，拟合程度较好。残差 ACF 和残差 PACF 在各延迟阶数都落在置信区间内，如图 3 所示，确定拟合后的残差项是白噪声序列。生成最终的预测结果，如图 4 所示。

表 3　模型统计量

模型	预测变量数	模型拟合度统计		Ljung-Box(18)			离群值数
		平稳 R 方	R 方	统计	DF	显著性	
ARIMA(1,1,0)	0	0.436	0.779		0		0

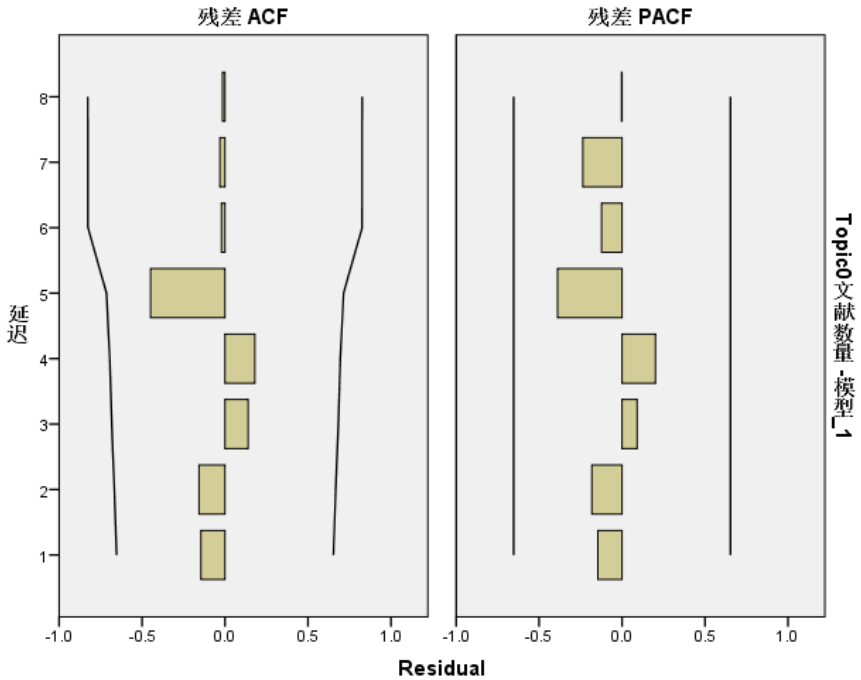


图 3　最终的残差 ACF 和残差 PACF 图

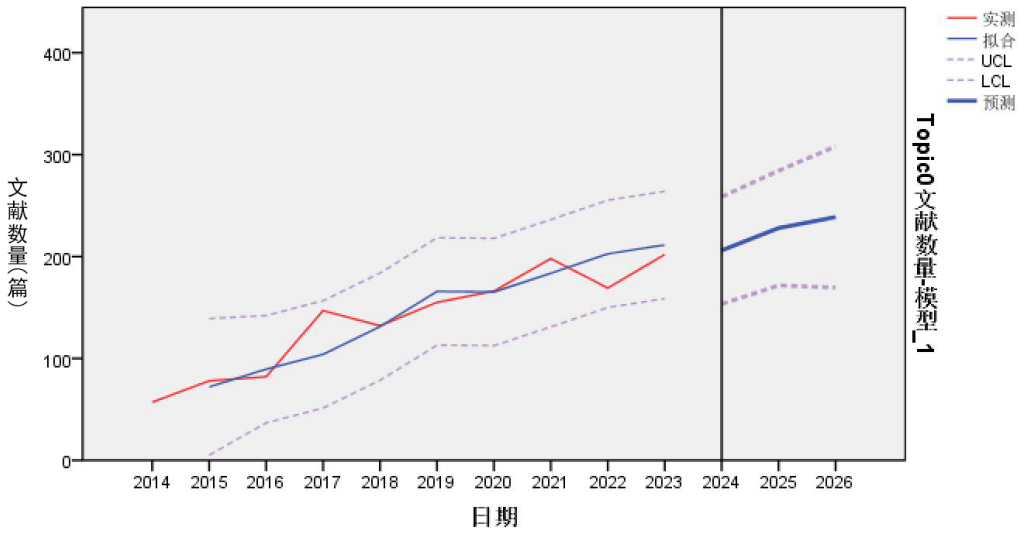


图 4　最终的数据预测图

采用上述方法，绘制全部 9 组折线图如图 5 所示。通过 9 组折线图，从文献数量上直观呈现出每组主题整体及包含的各 Topic 在 2014—2023 十年间的演变过程，以及后续可能的走向。

9 组主题随年份变化的发展趋势可以分为 4 种类型，分别是整体增长趋势、整体下降趋势、整体平稳趋势和异步趋势。

（1）整体增长趋势

Group1 电化学储能、Group2 可再生能源的催化反应、Group3 柔性可穿戴智能传感器系统、Group6 材料性能研究、Group8 太阳能电池五组主题处于整体增长趋势。

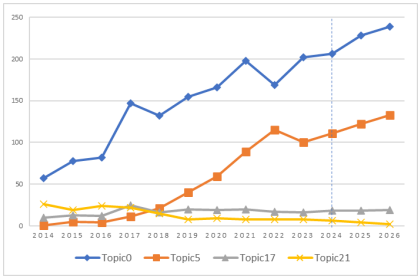
Group1 电化学储能主题组中，Topic0 和 Topic5 在十年间呈现出非常明显的上升趋势，尤其是 Topic5（锌离子电池），2024 年只有 1 篇高被引文献，到 2023 年单年产出量达 100 篇，在当年的 42 个主题中排名第三（前二分别是 Topic0 的 202 篇和 Topic1 的 172 篇）。科睿唯安发布的 2024 年研究前沿报告显示，锌离子电池的研究在热点前沿中排名 top7，并列入“重点热点前沿”，充分体现出该主题的重要性。

Group2 可再生能源的催化反应主题组中，Topic4 和 Topic1 的上升趋势最明显。这两个主题分别是关于光催化和电催化。根据预测，Topic4 和 Topic1 在 2024—2026 年间文献数量也处于明显的上升趋势。科睿唯安发布的 2024 年研究前沿报告显示，化学与材料科学领域 Top 10 热点前沿中有两项是关于催化主题的，体现出 Group2 的整体增长趋势与材料科学的热点前沿是契合的。

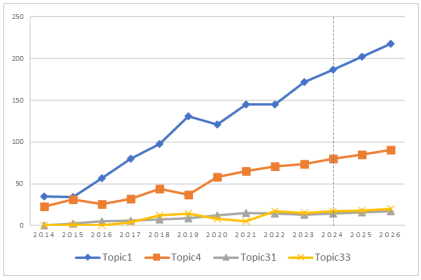
Group3 柔性可穿戴智能传感器系统主题组中，5 个主题都处于整体上升的趋势，且预测 2024—2026 年依然稳步上升。其中，Topic3 和 Topic9 近十年的涨势最明显，体现了水凝胶材料在柔性可穿戴智能传感器中的应用已是近年至未来一段时间的热门研究主题。

Group6 材料性能研究主题组中，Topic7 起点较高，增长幅度较缓；Topic13 曾在 2014—2017 年呈明显的下降趋势，后又重新回到增长趋势，并在 2022 年短时内超过 Topic7。这两个

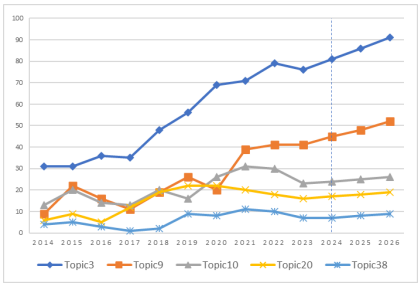
doi:10.3772/j.issn.2095-915x.2026.04.008



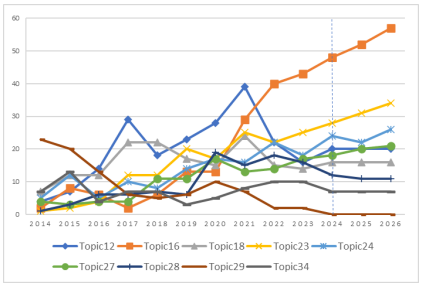
Group1 电化学储能



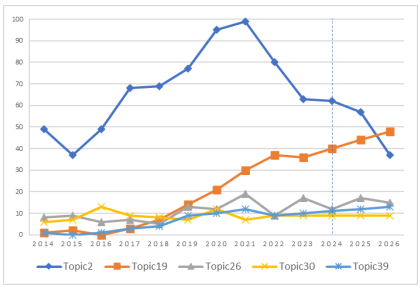
Group2 可再生能源的催化反应



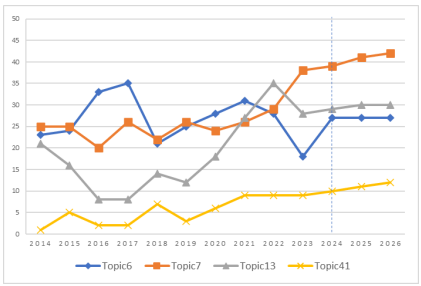
Group3 柔性可穿戴智能传感器系统



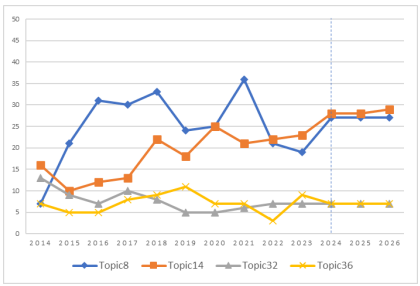
Group4 多功能复合材料



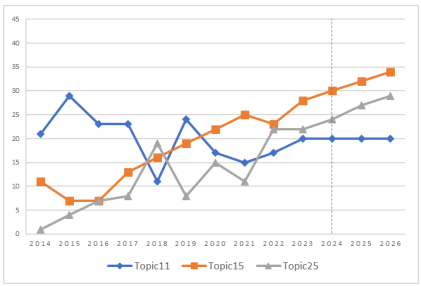
Grou5 生物材料的医学应用



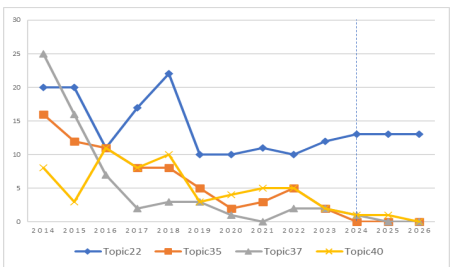
Group6 材料性能研究



Group7 光电器件中材料的选择和应用



Group8 太阳能电池



Group9 二维材料

图 5 9 组主题文献数量演化趋势

主题所研究的对象分别是高熵合金和混凝土，但都聚焦于材料的机械性能，一直都是研究的热点。

Group8 太阳能电池主题中，Topic11 趋势平稳，Topic15 和 Topic25 都呈现出上升的趋势。可以看出，对有机太阳能电池和钙钛矿太阳能电池的研究在未来一段时间内都将是比较热门的领域。

（2）整体平稳趋势

Group7（光电器件中材料的选择和应用）的 Topic8、Topic32、Topic36 在近十年的文献数量变化趋于平稳，只有 Topic14 有小幅的上升趋势。Group7 近十年的平稳趋势意味着这组主题的相关研究进入了相对成熟的时期 [22]，可能正面临新的理论突破，并产生更多新的研究主题（数据上可能体现为该主题的文献量下降）。

（3）整体下降趋势

呈现整体下降趋势的主题组只有 Group9 二维材料，除 Topic22 相对稳定外，其他几个主题都展示出明显的下降趋势。其中，Topic35 和 Topic37 这两个金属二硫化物相关的主题下降幅度最大，一方面说明金属二硫化物的研究在现阶段已经趋于完备，需要新的理论突破性研究；另一方面，这两个主题文献数量的逐年递减，也可能是相关研究进入成熟时期以后，产生了新的研究主题。如前文所述，新主题的增长体现在原主题上即是数量的减少。

（4）异步趋势

Group4 多功能复合材料、Group5 生物材料的医学应用 2 个主题组内的多个主题都出现了相反的趋势。

Group4 中的 Topic12、Topic18、Topic28、Topic34 十年间总体处于稳定趋势，Topic16、Topic23、Topic24、Topic27 持续上升，而 Topic29 显著下降，从组内文献数量最多主题降至最低。在储能、催化、传感、电磁屏蔽等领域，MXene 比石墨烯材料展现出更大的优势 [23]，有极大的应用前景。因此在文献数量方面，其他与能源存储、催化等相关的主题因为新技术的发展促进了高被引文献数量的上涨，同时围绕石墨烯展开研究的 Topic29 文献数量随之减少。

Group5 中，Topic19 在十年间从组内文献数量最少的每年 0 篇增至组内文献数量最多的每年 48 篇，涨幅较大；Topic2 从 2014 年的 49 篇，到 2021 年 99 篇增至最高点，之后持续下降，2023 年降至 37 篇，回到 2014 年相近的数量。因为 Topic2 十年间的数量变动显示出一种周期性变化的特点，故难以判断后续将趋于稳定还是继续下降，图中对 2024—2026 年的数量预测只能代表其中一种情况。因此，Group5 在未来一段时间内可能属于异步趋势，也可能属于整体上升趋势。

5 主要结论

本文通过 BERTopic 主题模型对 ESI 材料科学领域近十年的高被引文献进行了主题聚类分析。研究表明，近十年材料科学的研究热点主要分为 9 组，按照文献数量从高到低分别是①电化学储能、②可再生能源的催化反应、③柔性可穿戴智能传感器系统、④多功能复合材料、⑤生物材料的医学应用、⑥材料性能研究、⑦光电器件中材料的选择和应用、⑧太阳能电池、⑨二维材料。这 9 组热门研究方向，既可以作为材料科学领域学者、师生科研和论文选题方

doi:10.3772/j.issn.2095-915x.2026.04.008

向的参考，也可作为高校课程体系建设的依据。多数研究方向都体现出材料科学与其他学科的交叉融合，如 Group1、Group2、Group8 侧重于化学学科；Group3、Group5 则体现了材料科学和医学、生物学、计算机科学等学科的融合。近十年材料科学领域理工融合、医工融合的特点，正是我国“新工科”战略所重视的方向 [24]。学科的交叉融合更有利于推动学科的创新发展，促进综合性复杂性问题的解决 [25]。

对 BERTopic 主题模型的训练结果获取的 9 组主题（42 个 Topic）分别绘制了折线图，直观展示了各组主题近十年在文献数量上的演变过程，并使用 ARIMA 时间序列模型进一步预测了 2024—2026 年文献数量的趋势。Group1 电化学储能、Group2 可再生能源的催化反应、Group3 柔性可穿戴智能传感器系统、Group6 材料性能研究、Group8 太阳能电池五组主题处于整体增长趋势；Group7（光电器件中材料的选择和应用）在近十年的文献数量变化趋于平稳；Group9 二维材料呈现整体下降趋势；Group4 多功能复合材料、Group5 生物材料的医学应用 2 个主题组内的多个主题出现了相反的趋势。综合看来，材料科学领域在新能源转化与存储、医学器材等方面的研究在近期以及未来一段时间都是受到广泛关注的热门主题。根据科睿唯安发布的《2024 研究前沿》报告，碳量子点发光材料、太阳能电池和电催化剂等被列为化学与材料科学领域的未来研究前沿，如表 4 所示，可以验证本研究的分析结果，揭示材料科学领域未来具有潜力的主题。通过时间序列分析预测学科主题的未来演化趋势，并识别其在学科发展历程中的阶段定位，能够为研究人员把握学科动态规律、决策者科学配置科技资源提供重要参考。

表 4 《2024 研究前沿》化学与材料科学领域的 Top 10 热点前沿

序号	研究前沿	核心论文	被引频次	核心论文平均出版年
1	碳量子点发光材料	36	4092	2021
2	硒化锑薄膜太阳能电池	15	2241	2021
3	酸性析氧电催化剂	32	6134	2020
4	有机窄谱带发光材料	23	4243	2020
5	过渡金属催化的烯烃双官能团化反应	36	6370	2020
6	阴离子交换膜燃料电池和水电解池	26	6016	2020
7	吸附式大气集水研究	27	5174	2020
8	氧化还原液流电池	23	4118	2020
9	水系锌离子电池无枝晶锌金属负极设计	41	19689	2019
10	镧系单分子磁体	20	5090	2019

本文以 ESI 材料科学 2014—2023 年 Q1 期刊的高被引论文作为研究对象，偏向对高水平、高影响力论文的分析。除了高被引论文，以下几类研究成果在本研究中尚未囊括：（1）提出新兴理论，具有巨大潜力，但当前被引频次较低；（2）涉及多学科交叉的小众研究主题，未来可能解决关键技术难题，但因研究者较少等原因，文献数量和影响力都不突出。在下一步的研究中，可扩大研究范围，增强分析的全面性，及时发现潜力研究方向和新的学科发展点。

参 考 文 献

[1] 陈雅君 . 新工科背景下功能材料专业多元交叉复合型人才培养探索 [J]. 高教学刊 , 2024, 10(28): 63-66.

[2] 李红艳 . 高校 ESI 潜力学科发展态势及竞争力提升策略研究 —— 以河南工业大学材料科学为例 [J]. 江苏科技信息 , 2024, 41(12): 42-48.

[3] 王金丽 . 科技文献热点主题提取及演化研究 [D]. 昆明 : 昆明理工大学 , 2021.

[4] 黄东 . 近五年国际图书情报领域研究主题 —— 基于 WOS 的文献计量分析 [J]. 图书馆学刊 , 2021, 43(10): 98-104.

[5] WU H, YI H F, LI C. An integrated approach for detecting and quantifying the topic evolutions of patent technology: a case study on graphene field[J]. Scientometrics, 2021, 126(8): 6301-6321.

[6] 李俊炀 , 彭国超 , 张宁 . 图情领域虚拟现实研究热点主题和发展趋势分析 [J]. 图书馆建设 , 2023(4): 88-99.

[7] GROOTENDORST M. BERTopic:neural topic modeling with a class-based tf-idf procedure[EB/OL] [2025-04-20]. https://arxiv.org/abs/2203.05794

[8] LI Y, ZHANG C Y, ZHAO L F. Advances in Economics, Business and Management Research[M]. Atlantis Press International Bv, 2024.

[9] BANERJEE S, PAN A. From Colonial Legacies to Linguistic Inclusion: a BERTopic Enhanced Bibliometric Insight Into Global South Higher Education[J]. IEEE Access, 2024, 12: 117418-117435.

[10] 苗宏慧 , 全情爽 , 舒心 . 基于 BERTopic 的科技人才政策文本主题识别与量化分析 —— 以东北三省为例 [J]. 现代情报 , 2025, 45(4): 110-121.

[11] 张家惠 , 丁敬达 . 基于 BERTopic 和 LSTM 模型的新兴主题预测研究 [J]. 情报科学 , 2025, 43(1): 98-105.

[12] 杨思洛 , 于永浩 . 基于 BERTopic 模型的国内信息资源管理研究主题挖掘与演化分析 [J]. 情报科学 , 2024, 42(8): 12-21.

[13] 徐汉青 . 融合 BERTopic 和 LSTM 的 LIS 学科 AI 研究主题演变分析及趋势预测 [J]. 情报科学 , 2025, 43(5): 105-116.

[14] LAI F L, MIAO Y E, ZUO L Z, et al. Biomass-Derived Nitrogen-Doped Carbon Nanofiber Network: A Facile Template for Decoration of Ultrathin Nickel-Cobalt Layered Double Hydroxide Nanosheets as High-Performance Asymmetric Supercapacitor Electrode[J]. Small, 2016, 12: 3235-3244.

[15] LIAO H, GUO X L, WAN P B, et al. Conductive MXene Nanocomposite Organohydrogel for Flexible, Healable, Low-Temperature Tolerant Strain Sensors[J]. Advanced Functional Materials, 2019, 29(39): 1904507.

[16] CHENG W K, ZHU Y, JIANG G Y, et al. Sustainable cellulose and its derivatives for promising biomedical applications[J]. Progress in Materials Science, 2023(138): 101-152.

[17] JIN L, XIAO X, DENG W L, et al. Manipulating Relative Permittivity for High-Performance Wearable Triboelectric Nanogenerators[J]. Nano Letters, 2020(20): 6404-6411.

[18] CHEN J, HUANG X Y, SUN B, et al. Highly Thermally Conductive Yet Electrically Insulating Polymer/Boron Nitride Nanosheets Nanocomposite Films for Improved Thermal Management Capability. [J]. American Chemical Society Nano, 2019, 13(1): 337-345.

[19] ZHAO M Q, XIE X Q, CHANG E. Ren, et al. Hollow MXene Spheres and 3D Macroporous MXene Frameworks for Na-Ion Storage[J]. Advanced Materials, 2017, 29(37): 1702410.

[20] ZHOU C K, LIN H R, HE Q Q, et al. Low dimensional metal halide perovskites and hybrids[J]. Materials Science and Engineering R-Reports, 2019, 137: 38-65.

[21] ALAFEEF M, SRIVASTAVA I, ADITYA T, et al. Carbon Dots: From Synthesis to Unraveling the Fluorescence Mechanism[J]. Small, 2024, 20(4): 2303937.

[22] 邱均平 . 信息计量学 [M]. 武汉 : 武汉大学出版社 , 2007.

[23] 郑伟 , 杨莉 , 张培根 , 等 . 二维材料 MXene 的储能性能与应用 [J]. 材料导报 , 2018, 32(15): 2513-2537.

[24] 亓钧雷 , 从保强 , 宋晓国 , 等 . “新工科”背景下基于学科交叉的拔尖人才创新能力培养模式研究 [J]. 机械设计 , 2023, 40(1): 155-160.

[25] 张雪 , 张志强 , 朱冬亮 . 基于时间序列分析的潜在学科交叉前沿主题识别研究 [J]. 情报理论与实践 , 2024, 47(4): 152-162.

（责任编辑：李昶璇）