



开放科学
(资源服务)
标识码
(OSID)

多维评价视角下我国生成式人工智能服务关键共性技术识别研究

邵鹏 邱小洁

西安工程大学管理学院 西安 710048

摘要: [目的/意义] 生成式人工智能作为新一代信息技术的核心驱动力,正在深刻改变全球科技竞争格局。[方法/过程] 基于国家网信办发布的生成式人工智能服务备案信息,采用 LDA 模型识别专利技术主题,构建包括通用性、效益性、关联性、稳创性和重要性五个维度的关键技术主题量化指标,借助统计分析、社会网络分析方法对我国生成式人工智能服务关键共性技术进行识别。[结论/结果] 我国生成式人工智能服务包括 28 个技术主题,其中大模型训练与优化、用户对话与意图识别、语音文本识别与合成、图像生成与处理是我国生成式人工智能服务领域的关键共性技术主题。研究发现对我国生成式人工智能服务技术发展提供了参考依据,为关键共性技术的识别提供了研究方法。

关键词: 生成式人工智能服务; 大模型; 关键共性技术; 社会网络分析; LDA 主题模型

中图分类号: G306; G35

Research on the Identification of Key Generic Technologies in China's Generative Artificial Intelligence Services from a Multi-Dimensional Evaluation Perspective

SHAO Peng QIU Xiaojie

School of Management, Xi'an Polytechnic University, Xi'an 710048, China

Abstract: [Purpose/Significance] As a core driving force of the new generation of information technology, generative artificial intelligence is profoundly reshaping the global landscape of technological competition. [Method/Process] Based on the filing information of generative artificial intelligence services released by the Cyberspace Administration of China, this study applied the LDA model to identify patent technology topics. The paper constructed a quantitative indicator system for key technology topics, including five dimensions: generality, benefit, relevance, stability, and importance. Statistical analysis and social network

基金项目 陕西省教育厅科学研究计划青年创新团队项目“陕西省产业链创新联合体合作网络特征与政策促进机制研究”(24JP062); 西安工程大学学位与研究生教育综合改革研究与实践重点项目“生成式人工智能在研究生创新能力培养中的应用研究”(24yjzg10)。

作者简介 邵鹏(1987-), 通信作者, 博士, 副教授, 主要研究方向为创新管理与社会网络, E-mail: shaopengde@sohu.com; 邱小洁(2001-), 硕士研究生, 主要研究方向为创新网络分析。

引用格式 邵鹏, 邱小洁. 多维评价视角下我国生成式人工智能服务关键共性技术识别研究[J]. 情报工程, 2026, 12(1): 17-28.

analysis examined key generic technologies in China's generative artificial intelligence services. [Result/Conclusion] The study identifies 28 technology topics in China's generative artificial intelligence services, among which large model training and optimization, user dialogue and intent recognition, speech-text recognition and synthesis, and image generation and processing are identified as key generic technology topics. The findings provide a reference for the development of generative artificial intelligence service technologies in China and offer a research methodology for identifying key generic technologies.

Keywords: Generative Artificial Intelligence Services; Large Models; Key Generic Technologies; Social Network Analysis; LDA Topic Model

引言

党的二十届三中全会明确提出,要“加强关键共性技术、前沿引领技术、现代工程技术、颠覆性技术创新,加强新领域新赛道制度供给”。作为制造业转型升级的创新引擎,关键共性技术在促进产业格局提质增效、筑牢国民经济根基以及培育国际竞争力等领域具有不可替代的战略价值^[1]。伴随“互联网+”和工业4.0时代的到来,生成式人工智能蓬勃发展且加速迭代,已经在研发设计、生产制造等诸多领域崭露头角。生成式人工智能是指具备文本、图片、音频、视频等内容生成能力的模型及相关技术^[2]。作为支撑信息化转型、赋能数字经济发展的新型技术基座,生成式人工智能在提升国家战略高度与增强国际竞争力等方面发挥着举足轻重的作用。因此,精准识别与筛选关键共性技术已成为产业技术创新战略决策中亟待攻克的难题^[3],通过精准识别产业关键共性技术、剖析关键共性技术发展现状,能够为政府优化创新资源配置明确重点方向,并为企业攻克“卡脖子”难题提供决策基础^[4]。

专利文本作为专利文献的核心构成部分,对发明创新的技术内容以及法律保护范围进行

了详尽描述,是重要的技术情报来源。已有文献在智能网联^[5]、生物技术、信息技术产业^[1]等领域,开展了专利关键共性技术识别方面的研究,但对生成式人工智能服务关键共性技术开展的研究较少。基于此,本研究以生成式人工智能服务专利作为研究样本,遵循专利文献挖掘思路,借助LDA主题模型处理专利文本,挖掘并提取文本潜藏的技术主题以及对应的主题特征词。在此基础上,从多个维度对技术主题进行量化分析。通过技术主题强度和技术共现率评定技术主题“通用性”;依据专利申请量识别技术主题“效益性”;基于技术主题共现强度判断技术主题间的“关联性”;利用等级度和语料库标准差衡量技术主题的“稳创性”;借助社会网络分析方法计算技术主题的度数中心性、中介中心性和接近中心性,综合判断技术主题的“重要性”。通过对我国生成式人工智能服务产业关键共性技术进行识别,以期为生成式人工智能领域关键共性技术的研发规划与资源布局提供参考。

1 研究框架

为识别生成式人工智能服务的关键共性技

术，采用 LDA 主题模型、统计分析与社会网络分析等方法开展研究，研究框架如图 1 所示。首先，从专利数据库中提取专利摘要，进行分词和词性标注，然后通过停用词处理和 TF-IDF 文本向量化，为后续分析做准备。接着，利用 LDA 算法对文本进行主题建模，评估主题困惑

度和一致性，提取主题文档和特征词。最后，通过技术主题强度、共现率、专利申请量、技术主题共现强度、等级度、语料库主题标准差、网络中心性等指标，从通用性、效益性、关联性、稳创性和重要性五个方面对技术主题进行综合评估。

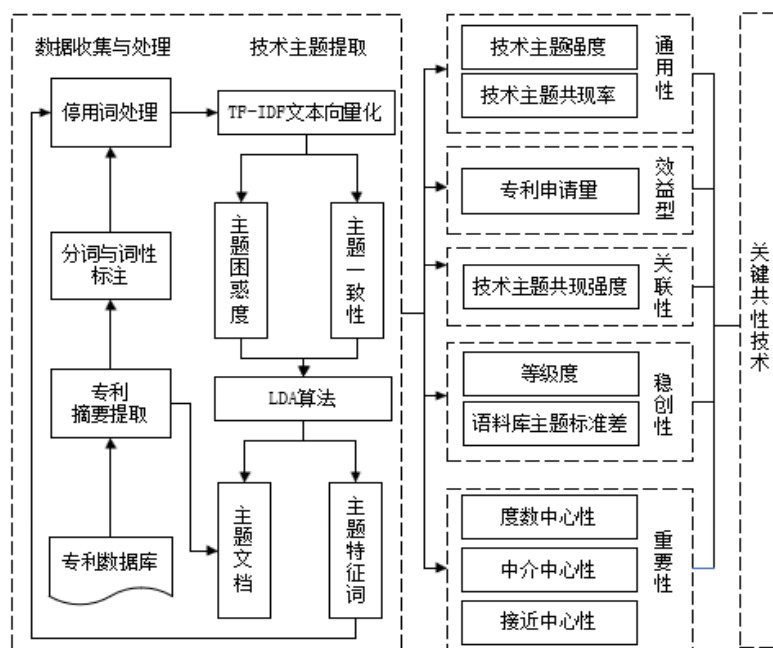


图 1 研究框架

基于关键共性技术五类特征，构建多维关键共性技术主题分析框架，如表 1 所示。结合不同主题对应的文档数量、主题概率技术强度、技术主题共现率、限制度等量化指标，设定合理阈值，筛选出候选的关键共性技术。

2 数据来源

截至 2024 年 12 月 31 日，共 302 款生成式人工智能服务在国家网信办完成备案。本文聚焦我国已备案主体的生成式人工智能服务领域

专利成果，以 256 家备案单位为基准，在大为专利数据库中制定检索表达式 TA=(生成 * 人工智能 or 通用人工智能 or AIGC or AGC or 生成式 AI or 大模型) 进行专利数据采集。考虑到 2017 年首次提出 Transformer 模型，随后 OpenAI 和 Google 相继推出了基于 Transformer 架构的 GPT-1 和 BERT 模型，本文将专利检索时间跨度设置为 2017—2024 年，检索共获得 5055 项专利数据。经过数据清洗，去除重复项及摘要为空的数据后，最终确定 3759 项有效专利文本作为生成式人工智能服务领域的研究样本。

表 1 关键共性技术主题分析指标

指标	衡量方法	计算方法	说明
通用性 ^[5, 7-8]	技术主题强度	$TS = \frac{\sum_{d=1}^{D_i} \theta_z^d}{D_i}$ (2)	TS 表示技术主题强度, θ_z^d 表示专利对于技术主题的隶属概率, D_i 表示时间段内样本总数, T 表示阈值。
	技术共现率	$T = \frac{\sum_d \sum_z \theta_z^d}{DZ}$ (3)	
	技术共现率	$count = \frac{count(Q(i,j) > e)}{h-1}$ (4)	$count(Q(i,j) > e)$ 代表主题 i 共现次数大于阈值 e 的主题数量, h 代表主题数。
效益性 ^[7]	专利申请数量	—	—
关联性	技术主题共现强度	$S(i,j) = \frac{coo(i,j)}{occ(i) + occ(j) - coo(i,j)}$ (5)	$S(i,j)$ 表示技术主题 i 与技术主题 j 的技术共现强度; $coo(i,j)$ 表示技术主题 i 与技术主题 j 的技术主题共现次数; $occ(i)$ 、 $occ(j)$ 分别表示技术主题 i 和 j 的出现频次。
稳创性 ^[8, 10]	语料库主题标准差	$CTS(i) = \sqrt{\frac{1}{M} \sum_{j=1}^M \sum_{i=1}^k (t_{ij} - \frac{1}{M} \sum_{j=1}^M t_{ij})^2}$ (6)	M 代表文档数量, K 代表主题数量, t_{ij} 表示主题 i 在第 j 篇文档中的概率。
	等级度	$H = \sum_j \left(\frac{C_{ij}}{C/N} \right) \ln \left(\frac{C_{ij}}{C/N} \right) / N \ln(N)$ (7)	C_{ij} 是限制度, N 是个体网络规模, C/N 为平均限制度, $N \ln(N)$ 则是可能的最大总和值。
重要性 ^[11-12]	度数中心性	$C_{D(i)} = \frac{deg(i)}{N-1}$ (8)	$deg(i)$ 表示与技术主题 i 直接相连的技术主题数量, N 表示网络中的技术主题总数。
	中介中心性	$C_{B(i)} = \sum_{s \in N} \sum_{t \in N} \frac{\sigma_{st}(i)}{\sigma_{st}}$, $s \neq t \neq i$ (9)	N 表示网络中的技术主题总数, σ_{st} 表示技术主题 s 与 t 之间存在的最短路径数量, $\sigma_{st}(i)$ 是技术主题 s 和 t 间存在的经过技术主题 i 的最短路径数量。
	接近中心性	$C_{C(i)} = \frac{1}{\sum_{j \in N} d_{ij}}$, $j \neq i$ (10)	N 表示网络中的技术主题总数, d_{ij} 表示技术主题 i 与网络中另一技术主题 j 之间的最短路径。

3 技术主题识别

3.1 技术主题抽取

本文依据困惑度和一致性确定最优主题数目。由图 2 可以看出, 困惑度曲线随着主题数目的增加呈下降趋势, 当主题数目为 28 个时, 困惑度曲线降至最低点; 由图 3 可以看出, 随着主题数目的增加, 一致性曲线不断上升, 当主题数目为 28 个时, 一致性曲线达到峰值。为了避免模型过拟合, 选择在困惑度值下降趋势、一致性值上升趋势均较为平稳时确定最佳主题数。本文基于 gensim 库

搭建 LDA 训练架构。将模型主题参数 Z 设置为 28, 参考已有研究, 设定参数 $\alpha=50/Z$, $\beta=0.01$ 。

3.2 文档-主题-特征词提取

将专利摘要分词结果存入 EXCEL 结果文件, 运用 Python 计算 LDA 主题模型, 提取并输出 28×20 的“主题-特征词”矩阵, 鉴于数据量较大仅对部分数据进行展示, 如表 2 所示。从技术主题及特征词提取结果看, 28 个生成式人工智能服务领域的技术主题均与生成式人工智能服务领域密切相关, 且各技术主题间具有显著差异性, 表明提取效果较为理想。

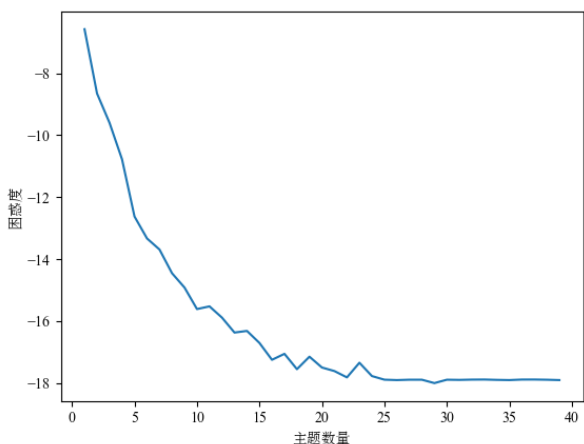


图2 困惑度图

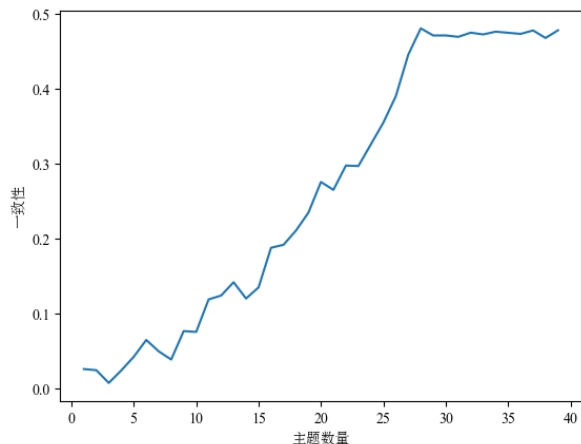


图3 一致性图

表2 技术主题及特征词(部分)

主题	主题关键词
Topic1	控制、状态、系统、调度、运行、电网、AGC、资源、功率、机组、决策、调频、实时、仿真、调节、优化、终端、频率、储能、传输
Topic2	参数、训练、计算、神经网络、预设、步骤、策略、迭代、执行、操作、大模型、优化、更新、阈值、评分、领域、量化、结构、性能、装置
Topic3	区域、位置、地图、道路、生成、兴趣、方向、要素、评测、定位、维度、点云数据、场景、装置、人工智能、点云、关联、技术领域、高精、信息
Topic4	图像、生成、人脸、图像处理、人工智能、装置、视觉、计算机、目标、原始、技术领域、获取、方案、风格、待处理、深度学习、区域、背景、电子设备、边缘
Topic5	样本、训练、生成、网络、方法、技术领域、深度学习、输入、装置、人工智能、目标、获取、预测、方案、调整、标签、对抗、参数、损失函数

一个文档隶属于多个主题，每个主题是一组单词及其概率。为排除低概率技术主题影响，通过 LDA 主题模型运算得到 3759×28 的“文档 - 技术主题概率”矩阵，

如表 3 所示。本文基于已有学者的研究^[7,14]，若主题分布概率大于 0.1，则判定该文档与该主题相关，进而删除文档中主题分布概率低于 0.1 的文档。

表3 文档及其技术主题概率(部分)

文档编号	主题	概率值	主题	概率值	主题	概率值	主题	概率值
1	1	0.244631	13	0.328168	23	0.135744	27	0.161797
2	1	0.592157	24	0.144052				
3	2	0.335906	15	0.574865				
4	2	0.538834	3	0.180389	16	0.209959		
5	3	0.535834	13	0.138698	14	0.144990		

4 关键共性技术分析

对 28 个技术主题依次开展通用性、效益性、

关联性、稳创性、重要性分析，从而挖掘出生成式人工智能服务领域的关键共性技术。

4.1 技术通用性分析

技术通用性分析是判断技术是否具有广泛应用潜力的重要依据。技术主题强度的高低直接关联其影响力大小，强度越高，表明该技术越具影响力，属于热门技术范畴。利用式(2)

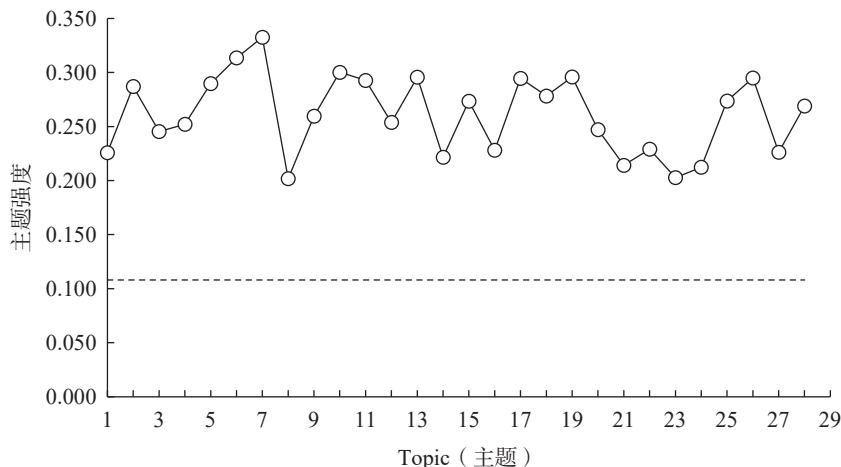


图4 技术主题强度及阈值

将矩阵导入 Gephi 中，绘制技术主题共现网络图，如图5所示。其中，各节点代表不同的技术主题，节点大小与技术主题在专利语料库中出现的频次高低相关，节点间的连线粗细象征着主题之间的关联强度^[16]。从网络图可以看出，Topic14、Topic17、Topic23 和 Topic28 节点相对较小，说明这四个主题在语料库中出

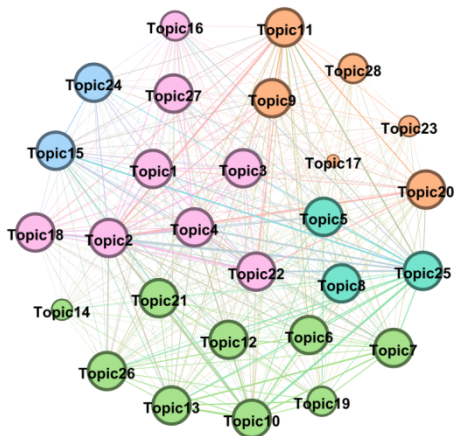


图5 技术主题共现网络图

计算出每个主题在技术领域的强度，再根据式(3)计算得阈值为0.108。从图4可以看出28个技术主题强度均高于阈值，进一步依据“文档-技术主题概率”分布矩阵，构建技术主题共现矩阵。

现的频次较低，影响力较弱，因此，将这四个主题予以剔除。

参考栾春娟^[9]的研究，设定共现频次阈值为100、技术共现率阈值为0.2。根据式(4)计算技术主题的共现伙伴数量及技术共现率(见表4)。结合“主题-特征词”矩阵和生成式人工智能领域的专利文献确定主题名称。经统计，共有16个技术主题的技术共现率高于0.2，其中Topic2(大模型训练与优化)、Topic4(图像生成与处理)、Topic10(虚拟形象与三维场景)、Topic25(语音文本识别与合成)等主题技术共现率达到0.5以上，Topic2(大模型训练与优化)技术共现率高达0.81，为所有主题中的最高值，这表明该技术的通用性最为突出，被广泛应用于自然语言处理、计算机视觉等领域。根据技术主题的技术共现率分析，初步筛选出通用性较高的技术主题。

表 4 技术主题共现伙伴数量、技术共现率及主题命名

技术主题	共现伙伴数量	技术共现率	主题名称	技术主题	共现伙伴数量	技术共现率	主题名称
Topic1	9	0.33	电网调度与功率控制	Topic12	8	0.30	信息抽取与事件生成
Topic2	22	0.81	大模型训练与优化	Topic13	10	0.37	图像识别与检测
Topic3	7	0.26	地图与定位技术	Topic15	11	0.41	向量矩阵与文本处理
Topic4	14	0.52	图像生成与处理	Topic18	6	0.22	自然语言查询与检索
Topic7	8	0.30	知识图谱与文档处理	Topic19	8	0.30	服务推荐与请求处理
Topic9	6	0.22	资源生成与配置	Topic22	13	0.48	用户对话与意图识别
Topic10	14	0.52	虚拟形象与三维场景	Topic25	23	0.85	语音文本识别与合成
Topic11	10	0.37	界面设计与交互	Topic26	6	0.22	图像特征提取与处理

4.2 技术效益性分析

对于能够为企业和社会带来巨大效益的技术，企业往往采取相应的专利保护措施^[17]。马永红等^[7]认为专利申请数量变化可以反映判断技术主题的关注度及经济价值。因此本文从技术主题专利申请量的逐年变化判断技术主题价值，进而分析其效益性，结果如图 6 所示。与其他技术主题相比，Topic9、Topic18、Topic12

整体的专利申请量较低，在后续的研究中，将不再考虑这 3 个技术主题。Topic19 在 2017—2022 年专利申请数量呈现上升趋势，但 2022 年之后专利申请数量不断下降，这一趋势表明该技术在演化过程中技术活跃度逐渐降低，由于市场或技术环境的变化而逐渐失去关注焦点，导致研究进展处于停滞状态。因此，后续分析中不再考虑该技术主题。

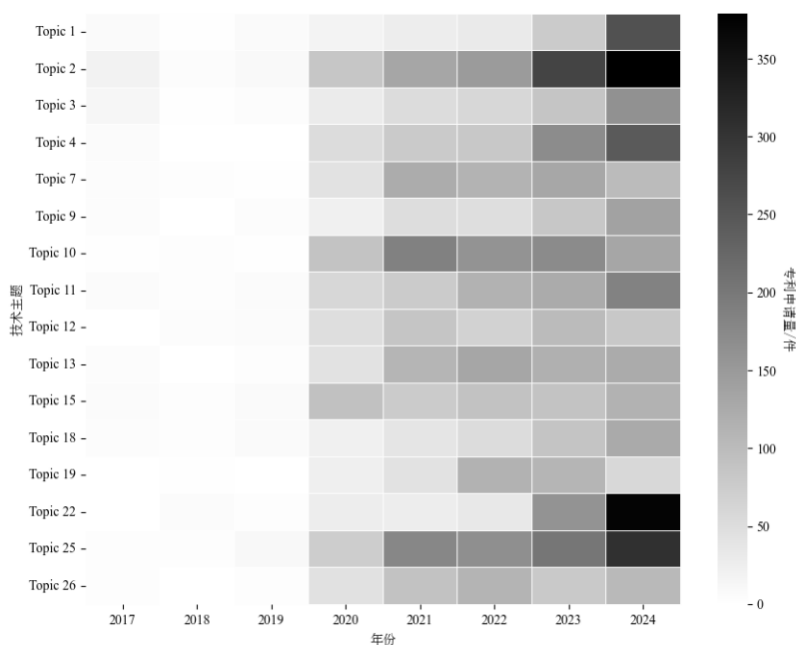


图 6 技术主题专利申请量变化趋势

4.3 技术关联性分析

产业关键共性技术的成果通常涉及多个学科的相关知识并包含多个产业部门的关联技术，故具有较强的关联性^[18]。经过通用性和效益性

分析后，还剩下 12 个技术主题。运用式（5）构建 12 个技术主题间共现强度的 Jaccard 系数矩阵，如表 5 所示，衡量技术主题间的潜在关联与强度等级。

表 5 技术主题共现强度的 Jaccard 系数矩阵(部分)

	Topic1	Topic2	Topic3	Topic4	Topic7	Topic10	……
Topic1	1	0.509	0.716	0.457	0.093	0.091	
Topic2	0.509	1	0.501	0.670	0.187	0.388	
Topic3	0.716	0.501	1	0.606	0.118	0.176	
Topic4	0.457	0.670	0.606	1	0.335	0.368	
Topic7	0.093	0.187	0.118	0.335	1	0.404	
Topic10	0.091	0.388	0.176	0.368	0.404	1	
……							

通过对每个技术主题与其他 11 个技术主题的共现强度进行加和汇总，得出每个技术主题的共现强度之和，结果如表 6 所示。可以看出，这 12 个

技术主题的共现强度之和均大于 2，表明这 12 个技术主题之间存在较高的共现强度，技术主题的彼此关联性较强，符合关键共性技术的基本特征。

表 6 技术共现强度

主题序号	主题名称	技术主题共现强度之和	主题序号	主题名称	技术主题共现强度之和
Topic1	电网调度与功率控制	3.791	Topic11	界面设计与交互	2.907
Topic2	大模型训练与优化	4.694	Topic13	图像识别与检测	3.244
Topic3	地图与定位技术	3.663	Topic15	向量矩阵与文本处理	2.980
Topic4	图像生成与处理	4.628	Topic22	用户对话与意图识别	4.676
Topic7	知识图谱与文档处理	2.657	Topic25	语音文本识别与合成	4.627
Topic10	虚拟形象与三维场景	3.647	Topic26	图像特征提取与处理	2.332

图 7 展示了 12 个技术主题的共现强度分布，整体以斜对角线为中心呈现对称分布。从图中可以看出，技术主题间关联性较强，尤其是 Topic2（大模型训练与优化）技术主题共现强度之和高达 4.694，说明该技术主题与其他技术主题在专利文本中共同出现的整体频繁程度较高，在生成式人工智能相关研究或应用中与多种技术主题存在较为紧密的关联；此外，Topic26（图像特征提取与处理）与其他大部分

技术主题共现强度较低，但其与 Topic2（大模型训练与优化）、Topic13（图像识别与检测）的共现强度均超过 0.3，表明两两技术主题组合在多篇专利中频繁出现，从技术关联角度看，大模型在训练与优化过程中，需处理大量图像数据，图像特征提取与处理能为大模型提供更优质、更具代表性的图像数据特征，帮助大模型更好地学习和识别图像模式，提升模型性能，而图像识别与检测作为连贯的技术流程，依赖

于准确的图像特征提取与处理。因此，主题之间存在较强的共现关系。通过技术主题共现强

度分析，本文筛选的技术主题“关联性”较强，符合共性技术“关联性”的基本特征。

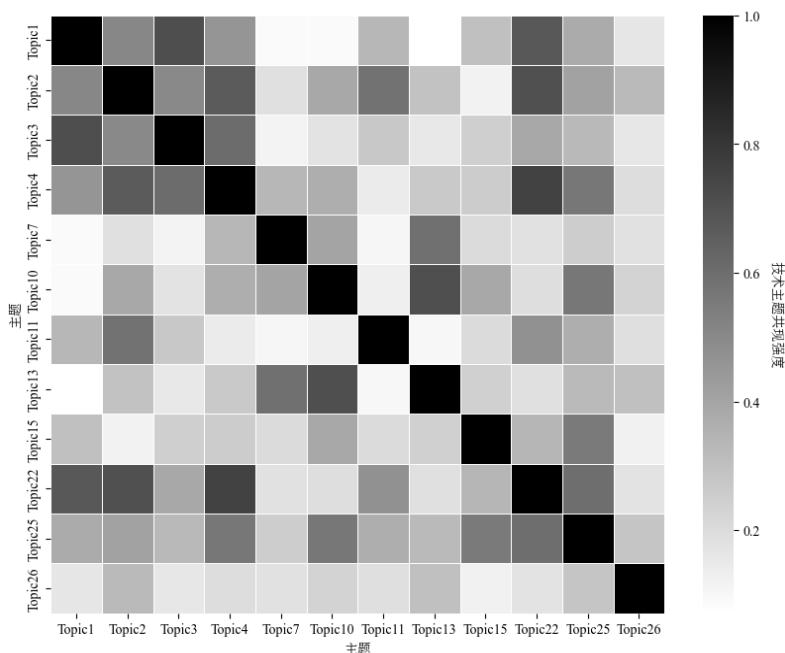


图7 技术主题共现强度

4.4 技术稳创性分析

借鉴徐宗煌等^[8]的研究，利用式(6)、式(7)计算生成式人工智能服务技术领

域专利主题创新性、稳定性指标，如表7所示，进一步明晰专利主题特征及发展潜力。

表7 技术主题稳创性

主题序号	H (创新性)	CTS (稳定性)	主题序号	H (创新性)	CTS (稳定性)
Topic1	-0.000199	0.116	Topic11	-0.000255	0.166
Topic2	-0.001397	0.166	Topic13	0.000199	0.165
Topic3	0.000492	0.128	Topic15	-0.000310	0.157
Topic4	-0.000030	0.134	Topic22	-0.000087	0.124
Topic7	0.000026	0.184	Topic25	-0.001201	0.156
Topic10	-0.000474	0.161	Topic26	-0.000199	0.162

创新性方面，用结构洞理论衍生的等级度指标衡量，可以量化技术主题在共现网络中利用结构洞获取异质信息的能力，进而反映其技术的创新性。该指标在不同专利主题间呈现正负差异与数值高低之分。正值主题表明相应技术主题在结构洞利用方面具备一定优势，拥有

更高的技术创新概率，如 Topic3 (地图与定位技术)、Topic7 (知识图谱与文档处理) 等。以地图与定位技术为例，随着智能出行、物流配送和共享经济等行业的迅猛发展，该技术已与多个领域建立了广泛而深入的联系。而负值主题则暗示这些技术主题在结构洞利用能力上

存在不足,技术创新的概率相对较低,如 Topic1(电网调度与功率控制)、Topic2(大模型训练与优化)等。以大模型训练与优化为例,由于大模型的结构复杂、参数庞大,模型训练推理带来的计算成本和能源消耗问题会成为掣肘领域发展的瓶颈之一^[19],使得该技术在跨领域创新方面进展相对缓慢。

稳定性方面,通过语料库主题标准差进行量化,反映了技术价值在不同文档中的波动情况。该值越高,意味着技术主题的概率分布越稳定,技术价值的可靠性越强。稳定性分布于0.116~0.184区间,体现出各专利主题技术价值稳定性的差异。例如,Topic1(电网调度与功率控制)的稳定性为0.116,处于较低水平,电网调度与功率控制技术主要应用于电力行业,由于该行业的专业性和特殊性,其技术体系相对封闭,与其他行业的技术交流和融合相对较少,在获取外部非重复异质信息方面存在困难,价值稳定性相对薄弱;Topic7(知识图谱与文档处理)的稳定性为0.184,相对较高,在数字

化信息爆炸的时代,各行业对知识管理和信息挖掘的需求与日俱增。知识图谱与文档处理可用于语义搜索、智能问答、个性化推荐等,在社交和电子商务等领域中实现价值^[20]。

4.5 技术重要性分析

栾春娟^[12]指出“多重测量中心度指标”较适合作为共性技术测度的网络中心性指标。若某共性技术主题在技术主题网络中具备高中介中心性及高度数中心性,则该主题可被视为关键共性技术主题^[21]。除考虑度数中心性、中介中心性指标外,本文还引入了节点的接近中心性^[11],以更精准地衡量技术主题在复杂网络中的影响力。将技术主题共现强度矩阵的每行均值设定为阈值并进行二值化处理,以技术主题作为节点,共现强度作为边构建网络。将处理后的数据导入Ucinet软件,利用式(8)至(10)计算各节点的度数中心性、中介中心性以及接近中心性,并对这些指标进行排名(表8)。结果显示,Topic2(大模型训练与优化),

表8 技术主题中心性

主题号	度数中心性		接近中心性		中介中心性	
	数值	排序	数值	排序	数值	排序
Topic1	59.259	6	71.053	5	1.467	7
Topic2	85.185	2	87.097	1	7.156	1
Topic3	55.556	8	69.231	7	0.993	10
Topic4	66.667	4	75.000	3	2.600	3
Topic7	48.148	10	65.854	9	0.736	12
Topic10	55.556	8	69.231	7	1.335	8
Topic11	55.556	8	69.231	7	2.079	5
Topic13	48.148	10	65.854	9	0.753	11
Topic15	59.259	6	71.053	5	2.169	4
Topic22	74.074	3	79.412	2	3.355	2
Topic25	88.889	1	90.000	1	6.544	1
Topic26	40.741	12	62.791	12	0.669	13

Topic22（用户对话与意图识别），Topic25（语音文本识别与合成）在三项指标中均排名前3，说明这些节点在技术网络中具有显著的影响力，处于核心地位；Topic4（图像生成与处理）虽然其度中心性处于第4位，但是其中介中心性以及接近中心性数值较高，因此本文认为其在技术网络中同样占据核心地位，对整体网络的控制力较强，与不同行业技术联系紧密，中介作用突出。

5 结语

本文基于专利文本构建关键核心技术识别模型，依托LDA主题模型与社会网络分析法，从通用性、效益性、关联性、稳创性与重要性五个方面识别了我国生成式人工智能服务领域的关键共性技术。在关键核心技术识别部分，研究通过计算困惑度和一致性确定了28个最优主题，在此基础上训练LDA主题模型，成功提取出相关技术主题以及“主题-特征词”矩阵和“文档-技术主题概率”矩阵，为后续研究提供了数据支撑。最后通过多维度指标组合分析，得出大模型训练与优化、用户对话与意图识别、语音文本识别与合成、图像生成与处理是当前生成式人工智能服务领域的关键共性技术。本文仍存在研究局限有待进一步研究。其一，本文数据来源于专利数据库，缺少论文、新闻、商业评论等多源数据的挖掘，数据覆盖面有待进一步拓展；其二，本文识别了生成式人工智能服务的关键共性技术，关于生成式人工智能关键共性技术演化与预测仍需进一步探索。

参考文献

- [1] 刘静, 胥彦玲, 张婧. 基于知识网络特征的产业关键共性技术识别模型与实证研究——以新一代信息技术产业为例[J]. 情报工程, 2024, 10(1): 28-41.
- [2] 卢宇, 余京蕾, 陈鹏鹤, 等. 生成式人工智能的教育应用与展望——以ChatGPT系统为例[J]. 中国远程教育, 2023, 43(4): 24-31.
- [3] XU Z. Machine learning-based quantitative structure-activity relationship and ADMET prediction models for ER α activity of anti-breast cancer drug candidates[J]. Wuhan University Journal of Natural Sciences, 2023, 28(3): 257-270.
- [4] PUC CETTI G, GIORDANO V, SPADA I, et al. Technology identification from patent texts: a novel named entity recognition method[J]. Technological Forecasting and Social Change, 2023, 28(3): 257-270.
- [5] 苏平, 李莹. 基于专利挖掘的智能网联汽车领域关键共性技术识别研究[J]. 创新科技, 2024, 24(12): 78-89.
- [6] 李阳, 唐玉洁. 基于专利信息的关键共性技术合作模式研究——以生物技术领域为例[J]. 科技进步与对策, 2023, 40(17): 67-77.
- [7] 马永红, 孔令凯, 林超然, 等. 基于专利挖掘的关键共性技术识别研究[J]. 情报学报, 2020, 39(10): 1093-1103.
- [8] 徐宗煌, 李小乐, 石进, 等. 基于多源异构数据的关键核心技术识别研究——以光刻技术为例[J]. 中国科技论坛, 2024(12): 127-136.
- [9] 栾春娟. 战略性新兴产业共性技术测度指标实证研究[J]. 中国科技论坛, 2012(6): 73-77.
- [10] BAI J. Panel data models with interactive fixed effects[J]. Econometrica, 2009, 77(4): 1229-1279.
- [11] 李瑞茜, 陈向东. 基于专利共类的关键技术识别及技术发展模式研究[J]. 情报学报, 2018, 37(5): 495-502.
- [12] 栾春娟. 网络中心性指标在技术测度中的应用[J].

- 科技进步与对策, 2013, 30(3): 10-13.
- [13] 缪青海, 王兴霞, 杨静, 等. 从基础智能到通用智能: 基于大模型的 GenAI 和 AGI 之现状与展望 [J]. 自动化学报, 2024, 50(4): 674-687.
- [14] 谭春辉, 熊梦媛. 基于 Ida 模型的国内外数据挖掘研究热点主题演化对比分析 [J]. 情报科学, 2021, 39(4): 174-185.
- [15] 吕文婷, 丁郡康. 新文科建设视角下我国档案学研究主题变化与未来展望——基于新文科建设前后档案学 cssci 期刊文献主题的比较分析 [J]. 档案管理, 2024(5): 5-10.
- [16] 齐虹, 闫静璇. 基于知识图谱的我国信息行为研究态势分析 [J]. 现代情报, 2018, 38(5): 131-139.
- [17] 黄鲁成, 张静. 基于专利分析的产业共性技术识别方法研究 [J]. 科学学与科学技术管理, 2014, 35(4): 80-86.
- [18] 郑彦宁, 浦墨, 刘志辉. 基于产业创新链的产业共性技术识别基本理论探讨 [J]. 情报理论与实践, 2016, 39(9): 53-58.
- [19] 车万翔, 窦志成, 冯岩松, 等. 大模型时代的自然语言处理: 挑战、机遇与发展 [J]. 中国科学: 信息科学, 2023, 53(9): 1645-1687.
- [20] 黄恒琪, 于娟, 廖晓, 等. 知识图谱研究综述 [J]. 计算机系统应用, 2019, 28(6): 1-12.
- [21] 栾春娟. 战略性新兴产业共性技术测度指标研究 [J]. 科学学与科学技术管理, 2012, 33(2): 11-16.

(责任编辑: 林毅)