



情报理论与实践

Information Studies: Theory & Application

ISSN 1000-7490, CN 11-1762/G3

《情报理论与实践》网络首发论文

题目： 数智融合视域下产业技术情报分析方法体系研究——基于图书情报领域论文的分析

作者： 安然，卢小宾，郑彦宁

网络首发日期： 2024-03-13

引用格式： 安然，卢小宾，郑彦宁. 数智融合视域下产业技术情报分析方法体系研究——基于图书情报领域论文的分析[J/OL]. 情报理论与实践.
<https://link.cnki.net/urlid/11.1762.G3.20240313.1318.003>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

●安 然^{1,2}, 卢小宾¹, 郑彦宁³

(1.中国人民大学信息资源管理学院,北京 100872;2.中国人民大学图书馆,北京 100872;3.中国科学技术信息研究所,北京 100038)

数智融合视域下产业技术情报分析方法体系研究*——基于图书情报领域论文的分析

摘要: [目的/意义]数智时代的到来深刻影响着产业技术情报分析范式与方法革新。面对数智环境的新变化和新要求,从情报分析方法角度加强大数据和大知识融合赋能产业技术发展。[方法/过程]基于内容分析法,通过分析论文的研究对象、数据资源、情报任务、产业领域、研究方法及其类别,深入探析研究方法、研究对象和情报任务的深层次关系,梳理数据驱动、知识驱动和融合驱动三种方法应用路径,全面反映和展示图书情报领域产业技术情报分析方法应用现状,构建数智融合视域下产业技术情报分析方法体系。[结果/结论]以数智融合的角度刻画产业技术情报分析方法现状以及应用情况,构建“数据—方法—任务”的产业技术情报分析方法体系。

关键词: 数智融合; 产业技术; 情报分析; 方法体系

Industrial Technology Information Analysis Method System in Perspective of Data Intelligence Integration: Based on the Analysis of Journal Literature in the Field of Library and Information Science

An Ran^{1,2}, Lu Xiaobin¹, Zheng Yanning³

(1.School of Information Resource Management, Renmin University of China, Beijing 100872; 2. Library, Renmin University of China, Beijing 100872; 3. Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: [Purpose/significance] With the arrival of the era of data intelligence, analysis pattern and methods of industrial technology information analysis have been changed. Faced with new changes and requirements of data intelligence environment, integrating big data and big knowledge in the perspective of information analysis promotes the development of industrial technology.[Method/process] Based on content analysis, the paper analyzes research objects, data resources, information analysis tasks, industrial fields, research methods and their classifications of journal literature in the field of LIS, explores the deep relationship among research methods, research objects and information analysis tasks and shows the methods' application paths which are divided into data driven, knowledge driven and fusion driven. The paper fully presents the current situation of application of industrial technology information analysis methods in LIS and constructs the industrial technology information method system in the perspective of data intelligence integration. [Result/conclusion] In perspective of data intelligence integration, the paper describes progress and application situation of industrial technology information analysis methods and constructs the industrial technology information analysis method system based on 'data-method-task'.

Keywords: data intelligence integration; industrial technology; information analysis; method system

0 引言

产业技术情报分析是提升产业竞争力的重要手段,力争以技术创新带动产业发展,在“未来技术”“未来产业”中占据主动权。产业技术情报分析历经从数据化、知识化、智能化到数智化的转变,每次外部环境变化都深刻影响着情报分析的范式与方法。自大数据纳入情报分析对象,情报链与数据链联系越来越紧密并互相融合。研究方法联合化是大数据环境下情报分析的一大特点,尤其是数据驱动和知识驱动的方法联合^[1],力图通过多种方法的组合、整合、融合抵消系统偏差,提升结论的准确性^[2]。因此,梳理数智融合视域下产业技术情报分析方法,构建产业技术情报分析方法体系对推动情报分析模式转变,促进产业技术情报分析实践发展具有重要意义。

1 数智融合对产业技术情报分析的影响

数智时代是“数据化”向“数智化”跨越^[3],给传统情报分析流程和服务模式带来了根本性改变,以大数据、人工智能为代表的情报分析方法是“一主三辅”情报工作模式中的重要一环^[4]。传统的情报分析方法已无法适应复杂异构的数据环境^[5],要求情报分析方法革新以满足

*本文为国家社会科学基金重点项目“新时期产业技术情报分析方法体系研究”(项目编号:21ATQ008)和国家重点研发计划课题“全球创新主体创新感知系统”(项目编号:2019YFA0707203)的成果。

下列需求：□应对小数据、大数据、小知识、大知识带来的挑战^[6]；□着眼外部多情境感知，由被动的情报处理分析转向主动的情报挖掘和知识服务^[7]；□促进多源数据、多维知识、数据知识的融合，强调广泛融合、深度融合和人机融合^[8]。数智融合环境下产业技术情报分析模式随之改变，嵌入越来越多数据融合和知识挖掘的模型和算法^[9]，注重基于智能技术和多源数据融合的情报分析^[10]。郭靖怡等基于规则和语法对专利数据进行文本挖掘，借助可视化方法展示产业技术链态势^[11]。Yun 等通过文本挖掘、聚类分析、模式识别，深入理解技术演进模式，对比目标领域和参考领域识别技术机会^[12]。Xu Haiyun 等利用链路预测和结构熵方法捕捉知识网络中结构变化用以预测突破性技术^[13]。Song 等采用结构方程模型分析调查问卷数据，评估用户对云计算技术的接受程度^[14]。刘明信等基于社会网络分析方法构建创新合作网络捕捉产业技术动态演化特征^[15]。为解决产业技术情报分析方法零散片段、体系化不强问题，霍朝光等从数据驱动视角构建了文本、网络、图像三维数据驱动的产业技术情报分析方法体系框架^[16]。

综上，产业技术情报分析方法从基于小数据、小知识点向大数据转变，强调数据驱动的分析思维，聚焦新技术、新模型和新算法的应用。但同时要防止“唯数据论”，忽略因果关系和揭示微观机制重要性^[17]，脱离产业技术本质规律，应该把握数据和知识双驱动的趋势，注重大数据和知识在研究方法层面的运用与融合^[18]。产业技术情报分析是图情领域（以下简称图情领域）研究的重要内容之一。数智环境正深刻影响图情领域的问题域、资源观等研究内容^[19]。现有情报分析方法体系构建多依靠主观判断、定性研究，无法更大范围和更细粒度展示研究方法全貌和梳理方法间关系^[20]。因此，本文从图情领域产业技术情报分析论文入手，通过内容分析，全面反映和展示图情领域产业技术情报分析方法应用现状，构建数智融合视域下产业技术情报分析方法体系。在理论层面，文章刻画了数智融合角度下产业技术情报分析方法现状以及应用情况，为零散碎片化的情报分析方法提供系统化架构，探索数智融合在产业技术情报分析方法中的应用路径和赋能方式。在实践层面，为情报工作人员在数智融合环境下开展产业技术情报分析提供借鉴参考，推动情报分析聚焦数据、知识和智能的融合，提升情报分析结论的科学有效性。

2 研究设计与方法

2.1 数据收集和处理

为全面了解国内外图情领域核心期刊上产业技术情报分析方法应用情况，同时选取 Web of Science 中“Information Science & Library Science”类别中 JCI>2（JCI 是跨学科衡量引文影响的指标，大于 1 代表高于平均水平）期刊的论文和中国知网中“图书情报与数字图书馆”类别中 CSSCI 索引的论文作为研究对象。考虑到查全率和查准率，检索策略将研究对象（产业技术、颠覆性技术、新兴技术、突破性技术、卡脖子技术、共性技术、持续性技术）和情报任务（情报分析、预测、识别、态势、创新、评估）交叉组配。

检索到的论文进一步人工识别筛选和向后引文补充。人工识别筛选的标准如下：□研究对象为产业技术；□研究目的为情报分析；□研究方法明确。排除标准：□综述性论文；□非研究性论文，如会议综述、导语等；□非产业层面论文，如企业内部技术创新和应用、国家政策分析等。文献检索时间为 2023 年 8 月 31 日，得到中文论文 137 篇，英文论文 56 篇，数据收集和处理流程如图 1 所示。

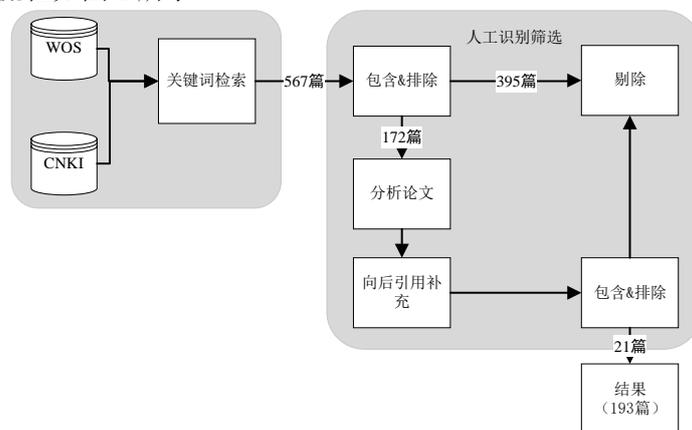


图 1 数据收集和处理

Fig.1 Data collection and processing

2.2 内容分析编码表构建

以每篇论文为分析对象,根据文献调研和小组讨论,确定类目编码体系表,每个维度均有明确判断标准,如表 1 所示。编码阶段由两名作者独立进行编码,并对编码结果进行汇总、对比和讨论。如存在争议,则咨询专家,直到达成一致。

表 1 类目编码体系表

Tab.1 Category encoding scheme

体系维度	编码	类目	判断标准和标注信息
文献时间	T1-T19	2005—2023 年	根据发表时间标注。
研究对象	A1	颠覆性技术	根据文章提及的研究对象技术类型进行标注,若未说明则标注 A7。
	A2	突破性技术	
	A3	新兴技术	
	A4	卡脖子技术	
	A5	共性技术	
	A6	特定技术	
	A7	未说明	
研究方法	B1	指标评价体系(层次分析法、模糊综合评价法、孔多塞评价法、灰色关联度分析)	根据文章提及的研究方法进行标注,一篇论文涉及一个或多个研究方法。
	B2	知识图谱、专利地图	
	B3	专家意见法(德尔菲法、访谈法、扎根理论、专家打分法、头脑风暴法)	
	B4	调查法(问卷调查法、网络调研法、案例分析法)	
	B5	统计分析法(描述性统计、聚类分析法、因子分析法、相似度分析、曲线拟合法、投入产出法)	
	B6	因果分析法(回归分析、结构方程模型)	
	B7	人工智能(机器学习、神经网络)	
	B8	系统分析法(产业技术路线图、内容分析法、DEMATEL 法、TRIZ、关联树)	
	B9	社会网络分析法(链路预测、PageRank 算法)	
	B10	特征指标法	
	B11	科学计量法(文献计量、替代计量、引文分析法)	
	B12	数据挖掘(文本挖掘、信息抽取、关联规则挖掘、模式挖掘)	

	B13	异常值探测法（突变检测、离群值检测）	
	B14	实验法	
	B15	计量经济分析法（CAPM、事件研究法）	
数据资源	C1	论文数据	根据文章提及的数据资源进行标注，若未涉及任何数据资源则标注为 C8。一篇论文涉及一个或多个数据资源。
	C2	专利数据	
	C3	统计数据	
	C4	访谈问卷数据	
	C5	项目成果数据	
	C6	媒体数据	
	C7	报告数据	
研究方法类别	D1	数据驱动	根据文中研究方法处理的数据资源类型进行标注。面向原始数据资源方法标注为 D1。面向结构化知识的方法标注为 D2。若同时面向数据和知识标注为 D3。
	D2	知识驱动	
	D3	数据和知识融合	
情报任务	E1	态势分析	根据文章要解决的情报任务进行标注，一篇论文涉及一个或多个情报任务。
	E2	识别分析	
	E3	预测分析	
	E4	评估分析	
	E5	创新分析	
产业领域	F1	电子通信领域	根据文章提及的产业领域进行标注，若未说明则标注 F12，一篇论文涉及一个或多个产业领域。
	F2	信息技术领域	
	F3	健康医疗领域	
	F4	制造领域	
	F5	能源领域	
	F6	材料领域	
	F7	生物技术领域	
	F8	农业领域	
	F9	军工领域	
	F10	环境领域	
	F11	金融服务领域	
	F12	未说明	

数据驱动和知识驱动的情报分析方法与数据资源类型息息相关。情报分析中通常不会明确区分数据或知识，可能是数据、信息、知识等各类信息形态的混合体。本文界定面向结构化、半结构化和非结构化原始数据资源开展情报分析，目的是挖掘数据间关联关系，在此过程中应用到的方法定义为数据驱动情报分析方法。数据驱动情报分析方法面对的数据资源有文本数据、网络数据、音频数据和图像数据等。面向结构化知识开展情报分析，包括专家知识和在数据资源基础上通过算法模型标注、组织、计算等产生的数据智能，在此过程中应用到的方法定义为知识驱动情报分析方法。知识驱动情报分析方法面对的数据资源有知识库、知识图谱、知识网络和专家意见等^[21-22]。

3 结果分析

立足研究方法角度，根据类目编码体系对各类目进行统计分析和关联分析。

3.1 描述性统计

图 2 展示了论文时间变化趋势,可以看出 2005 年产业技术情报分析就开始被关注,之

后缓慢上升，直到 2018 年开始发文数量有质的飞跃，随后一直保持高位状态。

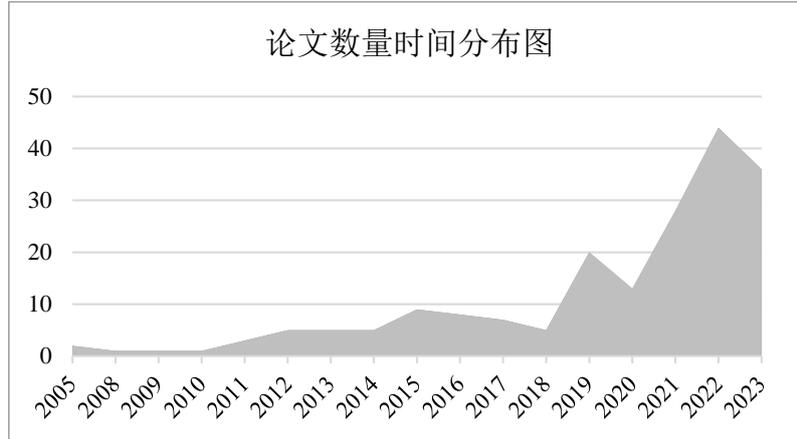


图 2 论文数量时间分布图

Fig.2 Time & papers' number distribution chart

各类目频次分布如表 2 所示。研究对象方面，6 类产业技术中研究最多的是新兴技术，其次是颠覆性技术，共性技术、突破性技术和卡脖子技术相对研究不多，没有持续性技术的研究。国内聚焦分析 6 类产业技术，6 类技术占比 58%，国外关注于特定技术的分析，特定产业技术占比 73%。国外研究角度更微观，研究问题更具体，倾向于分析具体技术。研究方法方面，使用频次最高的方法由高到低前 5 名分别是社会网络分析法、数据挖掘法、特征指标法、系统分析法、统计分析法。研究方法大多是组合使用，平均一篇文章使用两种分析方法。数据资源方面，由于专利文献可以全面反映科学技术发展动向^[23]，因此成为产业技术情报分析使用最多的数据资源类型。在科学大数据背景下，情报分析越来越重视多源数据的融合分析，因此数据资源呈现多样性的特点，平均一篇文章使用 1.2 种数据资源。研究方法类别方面，应用数据驱动研究方法最多，接下来是数据和知识融合研究方法以及知识驱动研究方法，越来越重视对数据和知识的应用和融合。情报任务方面，识别分析是最常见的研究任务，其次是态势分析和评估分析。国内外研究情报任务差异较大，国内研究聚焦于产业技术识别和预测分析，占比 64%；国外研究聚焦于产业技术效用、伦理、可行性、经济学等方面的评估分析，占比 59%。产业领域方面，产业技术情报分析应用的领域非常广，主要集中于信息技术领域、电子通信领域、健康医疗领域和制造领域。

表 2 类目频次分布

Tab.2 Category frequency distribution

体系维度	编码	类目	频次
研究对象	A1	颠覆性技术	35
	A2	突破性技术	6
	A3	新兴技术	38
	A4	卡脖子技术	6
	A5	共性技术	7
	A6	特定技术	47
	A7	未说明	54
研究方法	B1	指标评价体系	6
	B2	知识图谱、专利地图	12
	B3	专家意见法	27
	B4	调查法	30
	B5	统计分析法	39
	B6	因果分析法	14
	B7	人工智能	11
	B8	系统分析法	42
	B9	社会网络分析法	54
	B10	特征指标法	49

	B11	科学计量法	34
	B12	数据挖掘	53
	B13	异常值探测法	9
	B14	实验法	1
	B15	计量经济分析法	2
数据资源	C1	论文数据	39
	C2	专利数据	104
	C3	统计数据	9
	C4	访谈问卷数据	30
	C5	项目成果数据	3
	C6	媒体数据	4
	C7	报告数据	4
	C8	未涉及数据	26
研究方法类别	D1	数据驱动	73
	D2	知识驱动	54
	D3	数据和知识融合	66
情报任务	E1	态势分析	47
	E2	识别分析	72
	E3	预测分析	25
	E4	评估分析	42
	E5	创新分析	16
产业领域	F1	电子通信领域	28
	F2	信息技术领域	30
	F3	健康医疗领域	27
	F4	制造领域	21
	F5	能源领域	13
	F6	材料领域	7
	F7	生物技术领域	6
	F8	农业领域	4
	F9	军工领域	4
	F10	环境领域	6
	F11	金融服务领域	8
	F12	未说明	54

注：一篇文章可能涉及多种研究方法、数据资源、情报任务和产业领域。

3.2 研究方法趋势分析

产业技术情报分析应用的 15 类研究方法时间趋势如图 3 所示。从一开始产业技术情报分析就注重质性研究方法和量化研究方法的结合使用。但是 2017 年以前产业技术情报分析研究中应用到的量化研究方法大多基于抽样的样本数据等小数据。随着大数据技术在图情领域的广泛应用，2017 年开始，产业技术情报分析逐渐开始面向大数据应用机器学习、数据挖掘等方法。2020 年以后应用的研究方法主要是已有方法的深化，例如，在构建复杂网络基础上，应用链路预测的方法探测尚未连接但未来有较高可能产生连接的节点对，识别预测潜在新兴技术^[24]；在统计分析基础上，由于颠覆性技术、突破性技术的非线性生长，应用曲线拟合法绘制技术成长曲线预测未来技术走势^[25]。同时，也会借鉴应用其他学科领域的研究方法，利用突变值、离群值等异常值探测的方法识别颠覆性技术^[26]，利用眼动仪实验法探测消费者对技术过载压力的接受度^[27]等。

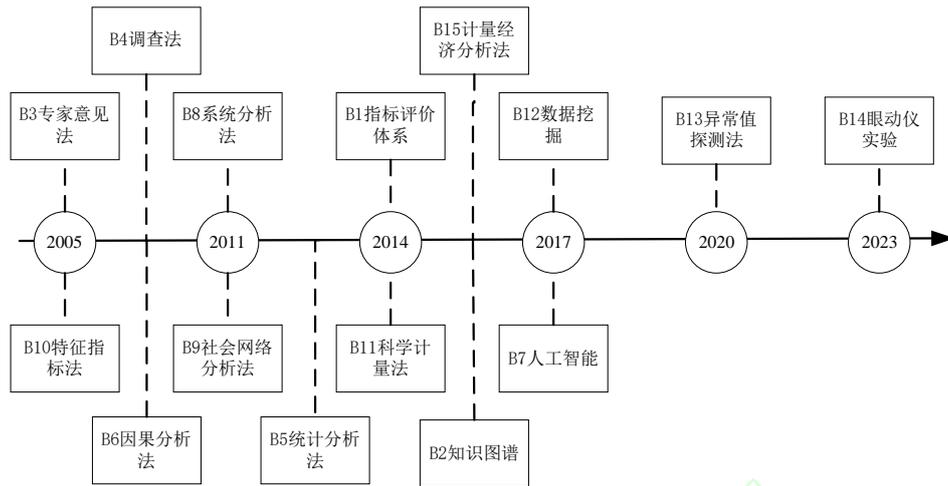


图3 产业技术情报分析方法时间演进

Fig.3 Time evolution of industrial technology information methods

3.3“研究对象—研究方法”关联分析

不同产业技术存在定义、内涵和特征等方面的差别，在技术创新、商业模式和国家安全战略各维度产生不同影响^[28]。构建“研究对象—研究方法”关联矩阵用于识别产业技术与情报分析方法选择的关系，如表3所示。颠覆性技术、突破性技术、新兴技术、卡脖子技术和共性技术5类技术在研究方法方面差异不大，社会网络分析、数据挖掘、特征指标和统计分析法都是常用的研究方法。除了常用的研究方法，5类技术在部分研究方法对应方面存在差异，相较于卡脖子技术和共性技术，颠覆性技术、突破性技术和新兴技术倾向利用科学计量法，基于现存知识，利用专利间引证关系寻找跨领域融合和知识重组过程中的技术机会^[29]；3类技术由于具备新颖性、异质性较高的特点^[30]，更多在社会网络分析的基础上采用突变理论和离群值等异常值探测法捕捉技术发展过程中的跃迁变化和范式革新。总的来说，目前不同产业技术概念边界模糊、交叉重叠，缺少各类技术概念的深入界定和辨析，导致概念混用或泛化的情况^[31]。因此，研究方法与研究对象的关联度不够显著。与5类技术相比，特定技术的情报分析通常聚焦于具体技术的具体分析，因此基于问卷调查的因果分析、基于专家意见和系统分析的理论研究以及基于文献计量和统计分析的现状分析较多。

表3 研究对象与研究方法关联矩阵

Tab.3 Incidence matrix of research objects and research methods

研究方法	颠覆性技术	突破性技术	新兴技术	卡脖子技术	共性技术	特定技术
B1 指标评价体系	1(1.3%)	0	3(3.4%)	0	0	0
B2 知识图谱、专利地图	1(1.3%)	1(6.2%)	1(1.1%)	0	0	1(1.1%)
B3 专家意见法	5(6.6%)	1(6.2%)	5(5.7%)	1(8.3%)	0	11(12.1%)
B4 调查法	7(9.2%)	0	1(1.1%)	0	0	22(24.2%)
B5 统计分析法	6(7.9%)	2(12.5%)	14(15.9%)	0	1(7.1%)	7(7.7%)
B6 因果分析法	0	0	0	0	0	12(13.2%)
B7 人工智能	1(1.3%)	0	7(8%)	2(16.7%)	0	1(1.1%)
B8 系统分析法	11(14.5%)	1(6.2%)	5(5.7%)	0	1(7.1%)	13(14.3%)
B9 社会网络分析法	12(15.8%)	3(18.8%)	12(13.6%)	2(16.7%)	5(35.7%)	5(5.5%)
B10 特征指标法	8(10.5%)	3(18.8%)	15(17%)	4(33.3%)	3(21.4%)	5(5.5%)
B11 科学计量法	8(10.5%)	3(18.8%)	6(6.8%)	0	0	7(7.7%)
B12 数据挖掘	12(15.8%)	1(6.2%)	17(19.3%)	3(25%)	4(28.6%)	3(3.3%)
B13 异常值探测法	4(5.3%)	1(6.2%)	2(2.3%)	0	0	1(1.1%)
B14 实验法	0	0	0	0	0	1(1.1%)
B15 计量经济分析法	0	0	0	0	0	2(2.2%)
总计	76	16	88	12	14	91

3.4“情报任务—研究方法”关联分析

情报任务是影响研究方法选择的重要因素，表 4 构建了“情报任务—研究方法”的关联矩阵。可以看出，情报任务和研究方法是强对应关系。识别分析常用数据挖掘方法、特征指标法和社会网络分析法。在识别分析前期的数据处理阶段，常基于主题模型（LDA、BERT、Word2vec 等）开展专利文本挖掘。在此基础上，应用社会网络分析法、特征指标法进一步遴选、识别产业技术。预测分析前期同样基于文本挖掘进行数据预处理，第二阶段多选用链路预测等社会网络分析方法和曲线拟合等统计分析方法开展预测分析。态势分析除经常应用数据挖掘方法以外，通过组合使用科学计量、描述性统计、社会网络分析法和知识图谱的方法揭示产业技术布局、演化和网络结构等现状和趋势。评估分析侧重于专家、市场对技术的效用、经济性、可行性、接受度等维度的反馈，通过访谈、专家打分等专家意见法和问卷调查等调查法纳入更多主观意见，配合回归分析、结构方程模型等因果分析法厘清影响技术发展的主要因素。技术创新是一个复杂多因素影响的过程，以创新系统、创新模式、TRIZ 等系统分析方法应用较多，运用决策试验和评价试验法(DEMATTEL)^[32]、预期技术伦理(ATE)^[33]等方法为系统分析产业技术各要素关系提供了途径。

表 4 情报任务与研究方法关联矩阵

Tab.4 Incidence matrix of information analysis tasks and research methods

研究方法	态势分析	识别分析	预测分析	评估分析	创新分析
B1 指标评价体系	1(1%)	3(1.7%)	0	2(2.3%)	0
B2 知识图谱、专利地图	8(8%)	2(1.2%)	0	2(2.3%)	1(2.2%)
B3 专家意见法	5(5%)	8(4.6%)	4(7%)	14(16.3%)	3(6.7%)
B4 调查法	9(9%)	3(1.7%)	1(1.8%)	20(23.3%)	3(6.7%)
B5 统计分析法	12(12%)	21(12.1%)	8(14%)	6(7%)	5(11.1%)
B6 因果分析法	1(1%)	0	0	11(12.8%)	2(4.4%)
B7 人工智能	1(1%)	4(2.3%)	4(7%)	1(1.2%)	0
B8 系统分析法	9(9%)	11(6.4%)	5(8.8%)	15(17.4%)	10(22.2%)
B9 社会网络分析法	14(14%)	30(17.3%)	13(22.8%)	2(2.3%)	7(15.6%)
B10 特征指标法	11(11%)	31(17.9%)	5(8.8%)	4(4.6%)	3(6.7%)
B11 科学计量法	13(13%)	11(6.4%)	3(5.3%)	3(3.5%)	7(15.6%)
B12 数据挖掘	15(15%)	44(25.4%)	10(17.5%)	3(3.5%)	4(8.9%)
B13 异常值探测法	1(1%)	5(2.9%)	4(7%)	0	0
B14 实验法	0	0	0	1(1.2%)	0
B15 计量经济分析法	0	0	0	2(2.3%)	0
总计	100	173	57	86	45

3.5 研究方法应用路径

管理决策发展历经“小数据”“小知识”“大数据”和“大知识”4 个阶段^[6]。根据内容分析，产业技术情报分析实现了从“小数据”向“大数据”的跨越，正在努力从“小知识”向“大知识”跨越。数据和知识融合实现了特征层面的初步融合。

3.5.1 数据驱动分析方法 产业技术情报分析的数据来源越来越广泛，由专利数据向论文数据、项目成果数据、媒体数据、统计数据等来源扩展和补充；数据规模越来越大，由百条数据规模发展到百万条数据规模；方法越来越智能，由人工统计向自动化算法转变。

数据驱动分析方法应用路径主要有三种，如图 4 所示。□基于“小数据”样本的分析和评价。采用抽样统计和概率分布等推论统计方法，从局部推断整体，具体包括描述性统计、因果分析、指标评价、实验等方法。这类方法关注局部样本中统计意义的显著性，探究数据间的深层次关系，但样本选择对结论的影响较大，无法面对大数据挑战。□基于数据特征关联的统计和分析。在多源数据计量和统计分析基础上，构建引文网络、合作网络揭示数据间的关联，构建量化测度指标揭示数据特征趋势，借助知识图谱、专利地图等可视化方法展示产业技术演化脉络、技术链图谱等。这类研究方法关注于数据外部特征，缺乏深入内容和语义层面的挖掘。□基于内容主题挖掘的识别和预测。通过分词技术、文本向量化、主题分类、语义聚类等方法挖掘文本内容，借助机器学习、人工智能实现大规模自动化挖掘，构建文本

语料库。在此基础上，组合使用知识图谱、特征指标、社会网络等方法识别和预测产业技术。这类方法实现大规模语义层面的数据挖掘，更能客观准确地识别和预测产业技术，但是存在结果可解释性弱、参数取值影响结果等问题。

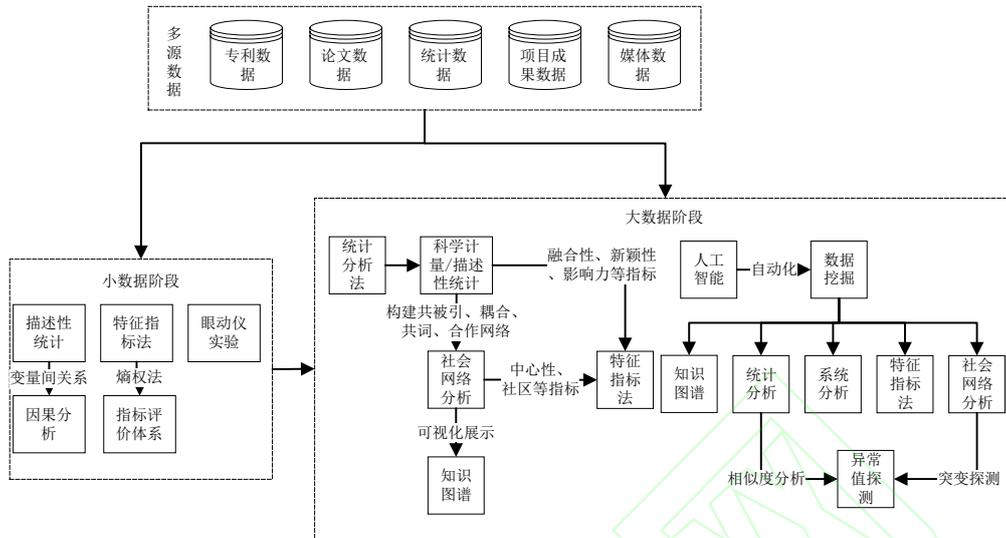


图4 数据驱动的分析方法应用路径

Fig.4 Application path of data driven methods

3.5.2 知识驱动分析方法 知识驱动分析方法主要基于专家知识和公众意见，由于缺乏产业技术情报领域的大型开源知识库和知识图谱，所以较少研究基于大规模数据智能。知识驱动的分析方法停滞在小知识阶段，难以实现向大知识跨越。

知识驱动分析方法应用路径主要有三种，如图 5 所示。□基于系统分析的创新和评估。概念框架是情报分析中数据信息解读的基础和条件^[34]。情报分析，尤其是创新分析和评估分析流程属性较强且受多因素影响，借助系统分析和模型思维，有利于明确分析目标和任务，规范流程，提高分析体系性和科学性。产业技术情报分析中已形成多种通用分析流程和模型，如产业技术路线图、TRIZ、产业链，或应用其他领域分析模型，如 DEMATEL、关联树法等，也有研究构建专用分析模型，如“创新漏斗”“产学研”等产业技术分析框架。系统分析法功能有限，难以开展理论性突破，新研究范式和方法会导致原有框架的失能。□基于专家知识的评估和预测。征求专家意见的方法有很多，包括德尔菲法、访谈法、头脑风暴法等。程序化的德尔菲法和访谈扎根分析适用于研究领域尚未形成理论框架的探索性研究^[35]。但是专家意见法主要依靠专家主观数据，成本较高，且无法大规模采集专家知识，缺乏对大数据的理解和应用。□基于公众意见的调查研究。调查研究的方法应用路径包括基于问卷调查的因果分析，从使用者角度开展产业技术的接受度、效用等评估；基于技术应用实例的观察，掌握产业技术的应用情况；开展国家、区域、产业、联盟等层面的网络信息调研掌握产业技术的发展状况。各类调研方法均受小数据的制约，不同样本数据可能颠覆结论。

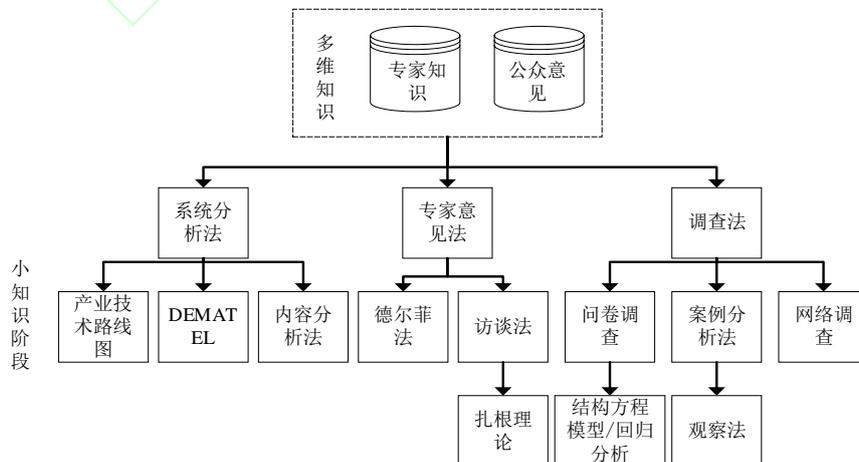


图5 知识驱动的分析方法应用路径

Fig.5 Application path of knowledge driven methods

3.5.3 数据和知识融合驱动分析方法 数据和知识是数据链和情报链融合关键环节,体现在特征层面、模型层面和决策层面的融合^[21]。产业技术情报分析中有意识地融合两类数据资源,但融合方法仍处在特征层面的探索,多为数据和知识粗犷地拼接、合并、相互验证,这类融合方法简单可行。

数据和知识融合驱动的分析方法应用路径主要有三种。□基于数据智能的分析。通过论文、专利等原始数据的组织、标注、关联等加工处理,构建语料库、知识图谱、知识网络等数据智能,开展基于数据智能的情报分析。如利用数据间引证关系、共现关系构建知识网络、知识图谱,提取网络中心势、聚集系数、网络规模等指标开展分析;利用自然语言处理技术开展信息抽取、语义表征、接近度计算等分析,构建语料库、语义矩阵等知识库用于进一步分析。这类研究方法实现了一定规模大数据向大知识的转变和分析,但是由于共享不足,局限于具体问题具体分析,没有形成成熟、开放的知识库。□基于专家意见的交叉验证。数据分析基础上,在技术清单梳理、检索式确定、结论识别判断等环节加入人工判读、专家判断、问卷调查等专家知识和公众意见进行交叉验证,提升结论的科学合理性。这类研究具备数据和知识融合的思维观念,但交叉验证缺乏统一流程和评价标准,效用无法考究。□基于质性研究的数据补充。在传统系统分析法、内容分析法、案例分析法、德尔菲法等质性研究方法基础上补充报告数据、统计数据等定量数据,实现理论研究和应用研究、内容分析和特征统计、技术预见和专利计量相结合。这类研究方法实现了定量分析方法和定性分析方法的相互补充,但仍然处于小数据和小知识的融合。

4 数智融合视域下产业技术情报分析方法体系构建

情报分析是多种要素、多维关系组成的系统,其中数据、方法和任务是相互联系的三大要素^[36]。基于上文分析,任务和数据影响方法,其中任务选择方法,数据决定完成方法的路径。因此,文章围绕“数据—方法—任务”三大情报研究的基础和本源,梳理数智视域下产业技术情报分析方法体系,如图6所示。

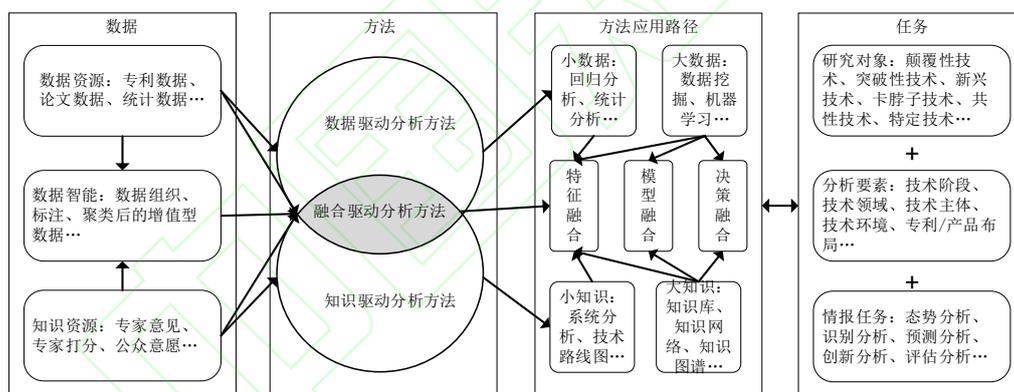


图6 数智融合视域下产业技术情报分析方法体系示意图

Fig.6 Industrial technology information analysis method system in perspective of data intelligence integration

数据方面,数智融合视域下产业技术情报分析方法应具备处理“数据+知识+智能”多源数据和多维知识融合的能力,提升产业技术情报分析方法处理“数据—信息—知识—智能”数智链条^[37]上高阶资源的能力。情报分析除基于专利数据、论文数据等数据资源以及专家知识和公众意见等知识资源,还要运用大数据、机器学习等技术将数据资源同知识资源相融合,通过数据组织、标注和揭示等方法构建一批提供开源数据智能的知识平台和知识库,实现小知向大知识的跨越。

方法方面,数智融合视域下产业技术情报分析方法在应用路径上呈现集成融合趋势,包括数据融合、知识融合以及数据和知识融合。目前很多研究正探索通过设计方法路径、优化模型算法等方式逐步向数据和知识融合方向发展,着眼于大数据和大知识在特征层面、模型层面、决策层面的融合方法。A.Smironov 等总结了知识积累、智能检索等 27 种知识融合模式,发现没有统一标准的知识融合方法框架,现有方法都是具体问题的具体分析^[38]。Cao Anye 等提出了一种基于注意力机制的矿山微震数据特征融合方法,实现了知识驱动和数据驱动在特征层面的融合^[39]。

任务方面,数智融合视域下产业技术情报分析方法应用最终目标是在复杂情境中面向各类技术完成具体任务,受研究对象、分析要素和情报任务共同影响。数智技术为完成不同类型情报任务提供了条件。为准确识别颠覆性技术,避免单一数据产生的偏差,Liu Xiwen 等采用“科学—技术—产业”链条上的多源数据,基于文本挖掘算法实现专家知识+多源数据的两阶段评估^[40]。数智技术也为在不同要素和角度上开展产业技术情报分析提供了可能。如 A.Bonaccorsi 利用社会网络分析方法构建英文版和德文版维基百科知识网络,并基于数据智能寻找德国新兴技术驱动因素^[41]。

5 总结与展望

数智时代的到来标志着从“数据化”向“数智化”的新跨越,深刻影响着产业技术情报分析范式与方法革新。通过内容分析方法构建编码表,分析国内外 193 篇图情领域产业技术情报分析论文的发表年份、研究对象、数据资源、情报任务、产业领域和研究方法及类别。立足于研究方法视角,通过“研究对象—研究方法”“情报任务—研究方法”的关联分析,发现情报分析方法与产业技术的关联度较弱,与情报任务的关联度较强。总结产业技术情报分析中数据驱动、知识驱动、数据和知识融合驱动三种方法应用思路,构建数智融合视域下产业技术情报分析方法体系。局限性:文章基于图情领域论文开展分析,数据具有一定的局限性。下一步将继续扩充数据的来源和范围,提高结论的适用性和可行性。

未来产业技术情报分析方法应从三个层面强化数智融合发展。□多维融合。实现“数据+知识+智能”的多源、多维、多模态资源的融合。□智能融合。利用大数据、人工智能、大模型等数智技术构建一批自动、快速、精准的数据智能库,推动基于数据智能的标准化、流程化、可复现的产业技术情报分析。□深度融合。探索小数据、小知识、大数据和大知识在模型层面和决策层面的深度融合,从简单拼接融合向更细粒度的融合发展。□

参考文献

- [1]罗立群,李广建.大数据环境下情报学发展的十个特征[J].图书与情报,2021(1):77-87.(LUO Liqun,LI Guangjian. Ten characteristics of the development of information science in big data environment[J]. Library & Information,2021(1):77-87.)
- [2]THOMSON M E,POLLOCK A C,ONKAL D, et al. Combining forecasts: performance and coherence[J].International Journal of Forecasting,2019,35(2):474-484.
- [3]陈国青,任明,卫强,等.数智赋能:信息系统研究的新跃迁[J].管理世界,2022,38(1):180-196.(CHEN Guoqing,REN Ming,WEI Qiang, et al. Data-intelligence empowerment: a new leap of information systems research[J]. Journal of Management World,2022,38(1):180-196.)
- [4]梁春华. 大数据与人工智能环境下“一主三辅”情报研究工作模式研究[J].情报理论与实践,2021,44(9):64-67.(LIANG Chunhua. “One domain and three aid” intelligence analysis model at big-data and artificial intelligence environment[J]. Information Studies: Theory & Application, 2021,44(9):64-67.)
- [5]申姝婧,杨建林.“数智赋能”及其背景下的情报思维培养[J].情报学报,2023,42(4):465-476.(SHEN Shujing,YANG Jianlin. Data intelligence empowerment and the cultivation of intelligence thinking under the background of data intelligence empowerment[J]. Journal of the China Society for Scientific and Technical Information,2023,42(4):465-476.)
- [6]MCDOWELL K. Storytelling wisdom: story, information, and DIKW[J]. Journal of the Association for Information Science and Technology, 2021, 72(10): 1223-1233.
- [7]MIHAELA M, SABIN M, RALUCA B, et al. Decision conceptual model for innovation ways using the competitive intelligence system[J].Ovidius University Annals: Economic Sciences Series, 2017, 17(1): 319-324.
- [8]李广建,罗立群.走向知识融合——大数据环境下情报学的发展趋势[J].中国图书馆学报,2020,46(6):26-40.(LI Guangjian,LUO Liqun. Towards knowledge fusion: the development trend of information science in big data environment[J]. Journal of Library Science in China, 2020,46(6):26-40.)
- [9]郑荣,王晓宇,高志豪,等.数智驱动背景下产业竞争情报智慧服务的认知框架与实现逻辑[J].情报学报,2023,42(7):761-774.(ZHENG Rong,WANG Xiaoyu,GAO Zhihao, et al. Cognitive framework and realization logic for industrial competitive intelligence service under digital intelligence background[J]. Journal of the China Society for Scientific and Technical Information,

2023,42(7):761-774.)

[10]安然,储继华,洪先锋.面向非结构化数据的情报分析方法体系框架研究[J].情报理论与实践,2024,47(2):143-150.(AN Ran,CHU Jihua,HONG Xianfeng. The method framework for unstructured data driven information analysis[J]. Information Studies: Theory & Application, 2024,47(2):143-150.)

[11]郭靖怡,王学昭,陈小莉.基于专利文本中产品关联关系的产业技术链构建与实证研究——以锂离子电池产业为例[J].图书情报工作,2023,67(5):108-118.(GUO Jingyi,WANG Xuezhao,CHEN Xiaoli. Construction and empirical study of industrial technology chain based on product association in patent text: taking lithium ion battery industry for example[J]. Library and Information Science, 2023,67(5):108-118.)

[12]YUN S,CHO W,KIM C, et al. Technological trend mining: identifying new technology opportunities using patent semantic analysis[J].Information Processing & Management,2022,59(4): 1-20.

[13]XU Haiyun, LUO Rui, WINNINK J, et al. A methodology for identifying breakthrough topics using structural entropy[J]. Information Processing & Management,2022,59(2): 1-20.

[14]SONG C H, KIM SW, SOHN Y W. Acceptance of public cloud storage services in South Korea: a multi-group analysis[J].International Journal of Information Management,2020(51): 1-12.

[15]刘明信,李丹丹,李荣,等.专利视阈下产业技术创新合作网络演化分析——以肿瘤疫苗为例[J].图书情报工作,2022,66(9):117-127.(LIU Mingxin,LI Dandan,LI Rong, et al. Evolution analysis of industrial technology innovation cooperation network from the perspective of patents: taking tumor vaccines as an example[J]. Library and Information Science, 2022,66(9):117-127.)

[16]霍朝光,卢小宾,杨冠灿,等.数据驱动的产业技术情报分析方法体系框架构建[J].图书情报知识,2022,39(1):73-83.(HUO Chaoguang,LU Xiaobin,YANG Guancan, et al. The method framework for data-driven information analysis towards industrial technology[J]. Documentation, Information & Knowledge, 2022,39(1):73-83.)

[17]黄晓斌,张明鑫.国外图书情报领域质性研究方法应用分析及启示——以 JASIS & T (1998-2018) 论文为例[J].情报科学,2021,39(2):96-105.(HUANG Xiaobin,ZHANG Mingxin. The analysis and implications on the qualitative methods used in the lis papers abroad: a case study of JASIS&T,1998-2018[J]. Information Science,2021,39(2):96-105.)

[18]卢小宾,安然.产业技术情报分析方法的现状与发展趋势[J].情报资料工作,2023,44(4):6-12.(LU Xiaobin, AN Ran. Analysis methods of industrial technology intelligence:status quo and development trend[J]. Information and Documentation Services, 2023,44(4):6-12.)

[19]孙建军,李阳,裴雷.“数智”赋能时代图情档变革之思考[J].图书情报知识,2020(3):22-27.(SUN Jianjun,LI Yang,PEI Lei. Some thoughts on the reform of library, information and archives management in the era of data intelligence empowerment[J]. Documentation, Information & Knowledge,2020(3):22-27.)

[20]李博闻,章成志.我国情报学研究方法体系构建研究的转向突破[J].情报理论与实践,2020,43(6):37-43.(LI Bowen, ZHANG Zhicheng. The turning and breakthrough in research paradigm of intelligence methodology system[J]. Information Studies: Theory & Application,2020,43(6):37-43.)

[21]卢小宾,霍帆帆,王壮等.数智时代的信息分析方法:数据驱动、知识驱动及融合驱动[J].中国图书馆学报,2024,50(1):1-18.(LU Xiaobin,HUO Fanfan, WANG Zhuang, et al. The methods of information analysis in data intelligence era: data driven, knowledge driven, and fusion driven by data and knowledge[J]. Journal of Library Science in China,2024,50(1):1-18.)

[22]邱韵霏,李春旺.智能情报分析模式:数据驱动型与知识驱动型[J].情报理论与实践,2020,43(2):28-34.(QIU Yunfei,LI Chunwang. Intelligent information analysis model: data-driven and knowledge driven[J]. Information Studies: Theory & Application, 2020,43(2):28-34.)

[23]孙红卫,田明.专利文献和技术创新[J].现代情报,2005(4):216-217.(SUN Hongwei,TIAN Ming. Patent documentation and technological innovation[J]. Journal of Modern Information, 2005(4):216-217.)

[24]王燕鹏,韩涛,陈芳.融合文献知识聚类和复杂网络的关键技术识别方法研究[J].图书情报

- 工作,2020,64(16):105-113.(WANG Yanpeng,HAN Tao,CHEN Fang. Identification of key technologies based on literature knowledge clustering and complex network[J]. Library and Information Science,2020,64(16):105-113.)
- [25]李树刚,刘颖,郑玲玲.基于专利挖掘的感知人工智能技术融合趋势分析[J].科技进步与对策,2019,36(23):28-35.(LI Shugang,LIU Ying,ZHENG Lingling. Analysis of technology integration trend of perceptual artificial intelligence based on patent mining[J]. Science & Technology Progress and Policy,2019,36(23):28-35.)
- [26]马铭,王超,张伟然,等.突变视角下潜在颠覆性技术识别与分析方法研究[J].情报理论与实践,2022,45(3):157-164.(MA Ming,WANG Chao,ZHANG Weiran, et al. Research on the methods of identifying and analyzing potential disruptive technologies from the perspective of catastrophe[J]. Information Studies: Theory & Application,2022,45(3):157-164.)
- [27]KIM S Y,PARK H,KIM H, et al. Technostress causes cognitive overload in high-stress people: eye tracking analysis in a virtual kiosk test[J]. Information Processing & Management,2022,59(6):1-15.
- [28]王康,陈悦,宋超,等.颠覆性技术:概念辨析与特征分析[J].科学学研究,2022,40(11):1937-1946.(WANG Kang,CHEN Yue,SONG Chao, et al. Disruptive technology: concept discrimination and characteristics analysis[J]. Studies in Science of Science, 2022,40(11):1937-1946.)
- [29]SCHOENMAKERS W,DUYSTERSA G. The technological origins of radical inventions[J].Research Policy,2010,39(8):1051-1059.
- [30]ROTOLO D,HICKS D,MARTIN B R. What is an emerging technology?[J].Research policy,2015,44(10): 1827-1843.
- [31]曲冠楠,陈凯华,陈劲.颠覆性技术创新:理论源起、整合框架与发展前瞻[J].科研管理,2023,44(9):1-9.(QU Guannan,CHEN Kaihua,CHEN jin. Disruptive technovation: origins, integrated framework, and prospects[J]. Science Research Management,2023,44(9):1-9.)
- [32]SHARMA M, LUTHRA S, JOSHI S, et al. Implementing challenges of artificial intelligence: Evidence from public manufacturing sector of an emerging economy[J]. Government Information Quarterly,2022,39(4): 1-14.
- [33]BREY PAE. Anticipating ethical issues in emerging IT[J].Ethics and Information Technology,2012,14(4):305-317.
- [34]HEUER JR RJ. Psychology of intelligence analysis[M]. Washington DC: United States Government Printing Office,1999:85-94.
- [35]URQUHART C,LEHMANN H,MYERS MD. Putting the ‘theory’ back into grounded theory: guidelines for grounded theory studies in information systems[J].Information Systems Journal,2010,20(4):357-381.
- [36]祝振媛.多维视角下的情报分析模型构建——从任务、数据、方法三维度出发[J].情报理论与实践,2021,44(5):36-43.(ZHU Zhenyuan. Research on information analysis model building from the three-dimensional perspectives of task, data and method[J]. Information Studies: Theory & Application,2021,44(5):36-43.)
- [37]杨建林,苗蕾,申姝婧,等.数智赋能科技情报服务能力现代化的动因与进路研究[J].科技情报研究,2023,5(4):1-16.(YANG Jianlin,MIAO Lei,SHEN Shujing, et al. Research on the motives and approaches for the modernization of technological intelligence service capability empowered by data and artificial intelligence technology[J]. Scientific Information Research,2023,5(4):1-16.)
- [38]SMIRNOV A, LEVASHOVA T. Knowledge fusion patterns: a survey[J]. Information Fusion,2019(52):31-40.
- [39]CAO Anye, LIU Yaoqi, YANG Xu, et al. FDNNet: knowledge and data fusion-driven deep neural network for coal burst prediction[J]. Sensors,2022,22(8):1-17.
- [40]LIU Xiwen, WANG Xuezhao, LYU Lucheng, et al. Identifying disruptive technologies by integrating multi-source data[J]. Scientometrics,2022,127(9):5325-5351.
- [41]BONACCORSI A, CHIARELLO F, FANTONI G, et al. Emerging technologies and industrial leadership. a wikipedia-based strategic analysis of industry 4.0[J].Expert Systems with Applications,2020(160): 1-18.

作者简介:安然,1993年生,博士生,馆员。卢小宾,男,1963年生,教授,博士生导师。郑彦宁(通信作者,Email:ynzheng@istic.ac.cn),男,1965年生,研究馆员。

录用日期:2024-03-05