

面向技术创新的企业智能情报服务平台构建

冯向梅¹ 顾方¹ 鲁瑛¹ 周群芳² 王俊松¹ 肖甲宏¹

(1. 中国化工信息中心有限公司情报资讯事业部, 北京 100029; 2. 宝武集团(宝钢)中央研究院, 上海 200126)

摘要: 本文设计并开发了反应快速、分析深入、功能集成、自由灵活、可视化展现的面向技术创新的“一站式”企业智能情报分析平台。该平台综合运用多源数据融合、文本聚类、关系发现等智能技术, 实现智能、高效、准确的技术信息获取, 不同主题领域技术识别与评估, 不同技术领域关系发现, 重点企业研发布局挖掘, 用户自助智能服务等功能。平台为企业用户提供领域技术的知识结构、发展趋势、优势资源等方面的分析, 并以多种可视化方式呈现高效的战略决策支持服务, 为企业技术情报研究业务提供有效便捷的研究方法和分析工具, 提高情报服务能力和效率。

关键词: 多源数据; 关系发现; 前沿技术; 技术预见; 信息生态

中图分类号: G255.51; G358 **DOI:** 10.3772/j.issn.1673-2286.2022.05.004

引文格式: 冯向梅, 顾方, 鲁瑛, 等. 面向技术创新的企业智能情报服务平台构建[J]. 数字图书馆论坛, 2022(5): 14-23.

鉴于日益复杂多变的信息环境和信息服务竞争不断加剧的状况, 加之面临贸易冲突升级的国际局面, 以及提升自主创新能力的迫切需求, 科技创新成为企业降本增效和增强核心竞争力的关键, 受到企业广泛重视。同时, 国家密集出台多项科技创新政策, 支持和鼓励企业科技创新战略方案的实施。

技术创新作为推动科技创新发展的关键举措, 不仅能够实现关键技术、核心技术、“卡脖子”技术攻关, 大力提升企业自主创新能力, 提升产业链现代化水平, 掌握竞争和发展的主动权, 还能够与竞争企业保持信息对称^[1], 尽快将技术成果转化为产业竞争优势。因此, 如何开展技术创新成为每个企业发展科技创新要思考的关键命题, 而提供全面、准确、及时、可靠的技术情报服务是开展技术创新的必要条件。

本文以需求为导向, 通过梳理目前国内外企业情报服务平台建设现状, 分析新形势下企业开展情报服务的需求, 以此研发面向技术创新的企业智能情报服务平台。通过海量异构数据的汇聚和整合, 基于多源数据

融合技术、文本挖掘技术、机器学习技术, 不仅实现领域技术识别、领域技术评估、领域技术关系挖掘等面向技术生命周期的一站式管理, 而且能够实现机构、专家等优势资源的多维分析, 从而为企业技术情报研究提供高效工具, 为企业技术布局提供战略决策支撑。

1 技术创新的概念界定

技术创新的思想可追溯于熊彼特的创新理论^[2], 之后众多学者从行为、经济、产品、时序等角度对技术创新的概念进行表述。从狭义角度讲, 技术创新是指人们在新产品开发或老产品改进中重新组织生产条件和要素, 创造性地运用不同的方法、工艺、工具或装备的过程, 并取得显著经济效益或具有潜在长远的经济效益, 通常包括产品创新和过程(工艺)创新等。从广义角度讲, 技术创新是指人们在生产实践活动中重新组织生产条件和要素, 创造性地运用其在科学实验和生产活动过程中所积累知识、经验和技能的过程, 并取得显著

的经济效益或具有潜在长远的经济效益,通常包括产品创新、过程(工艺)创新、市场创新、组织创新和制度创新等。

钱旭潮等^[3]认为企业技术创新经历了微创新、集成创新、延伸创新和持续创新阶段。微创新是零件层面,是指企业能够进行简单的单项技术创新,大大提高了该零部件的性能、质量,或生产效率。集成创新是产品层面,是指基于多项微创新,发明一个新产品或改进现有产品的核心技术。延伸创新是系统层面,是指在产品的基础上加入衍生产品或服务形成一个整体。持续创新是领域层面,是指企业能够继续保持领域或行业内的技术领先,主导或参与行业标准制定,是技术的领跑者。

因此,无论是从定义内涵还是从发展阶段来说,都需要对现有市场、产品、竞争对手所研发的技术情况进行全面系统的梳理和掌握,对世情、行情、市情有精准快速了解,从而为企业技术研发、技术创新和战略部署提供支撑依据。

2 企业技术情报服务相关研究

通过文献调研发现,目前企业技术情报服务相关研究主要围绕技术情报的服务方法、服务平台、服务内容3个方面展开。

技术情报服务方法研究主要采用以下3种方法:①以德尔夫法^[4-6]为主,侧重于国家级或重点领域的技术预见研究;②采用机器学习、文本挖掘等智能算法对技术情报进行分析,如许学国等^[7]采用机器学习方法对新能源汽车核心技术识别,周源等^[8]采用机器学习方法对机器人领域的新兴技术进行识别,王秀红等^[9]基于BERT-LDA对农业机器人的关键技术进行识别;③采用智能算法、评价模型等多种组合方法对技术情报进行分析,如TF-IDF和多指标评价模型^[10]、LDA主题模型和ROST^[11]、知识聚类 and 链路预测方法^[12]等。

技术情报服务平台研究主要集中在风电、电力、轨道交通等领域。鲁啸等^[13]以风电企业技术情报平台建设为例,针对资源利用不足、工作管理不便、沟通成本较大、成果共享范围窄等现存问题,设置动态信息、资料库、资源导航、情报需求、情报团队等栏目,用于服务内部技术创新。刘民^[14]充分利用元数据资源整合、知识发现、可视化图谱等技术,构建电力企业竞争情报服务

平台,提出SNS个人知识管理、综合趋势分析、机构内部评估分析、情报挖掘服务、可视化图谱、人才评估分析、调研辅助等创新服务。李子林等^[15]充分利用海量多源异构数据归集、基于智能算法的科技资源深度组织等技术,满足我国轨道交通行业科技情报服务存在整合海量多源异构数据、实现数据驱动科研热点研判、提供聚焦领域知识细分的专题定制以及搭建行业高端知识交流社群等发展需求。

技术情报服务内容研究主要围绕技术生命周期展开,具体包括以下3类。①不同技术类型的识别,如娄岩等^[16]利用专利文献和技术的关联关系进行替代性技术识别;吴菲菲等^[17]从创新性和学科交叉性角度构建前沿技术探测模型;黄鲁成等^[18]基于专利分析的客观分析方法为基础构建了颠覆性技术识别框架;此外,还有学者对新兴技术^[19-20]、前沿技术^[21]、核心技术^[22]、关键技术^[11, 23]等方面展开研究。②技术跟踪和监测,如崔怡雯等^[24]提出面向颠覆性创新的领域技术监测分类体系。③技术预见或技术预测,既有日本、韩国等开展的国家级技术预见,也有面向智能制造^[25]、生物医药^[26]、全球能源^[27]等不同科技领域的技术预测等。

由此看出,目前企业技术情报服务呈现以下特点:第一,大数据、文本挖掘、可视化等智能技术已成为技术情报服务平台建设的必要手段;第二,海量异构资源的整合成为技术情报服务平台建设的必要基础;第三,技术情报服务平台的研发、应用和推广还有待深入。虽然目前技术情报服务内容研究较为丰富,但技术情报服务平台(或系统)的研发并不多见,且大多涉及概念模型、系统架构、模块功能等设计层面,对技术识别的关键技术及技术关系揭示、技术类型评价等细节内容涉及较少,有待于领域技术识别、技术预见等模块和功能的进一步研发和应用推广。

3 面向技术创新的企业技术情报服务需求分析

经文献调研发现,企业技术情报服务在分析内容、分析算法、分析内容等方面已有较多研究成果,但仍存在分析结论不便于管理与共享、用户使用不够便捷、技术分析缺乏连续性等不足。因此,构建灵活、精准、智能、专业的企业技术情报服务平台成为企业持续技术创新的基础和必需。

3.1 基于多源数据、智能集成的平台结构

企业的业务场景复杂,且技术载体日益多元,为了避免决策结果偏差,保障结果全面准确和科学可信,需要海量多源异构数据作为支撑。针对海量异构的资源需求,传统人工上传的方式已不能满足数据采集的需要,因此需要借助智能工具实现高效的多源数据智能采集和集成,对平台结构方面的需求主要集中在以下3个方面。第一,多源数据的智能爬取。通过研发网络爬虫工具,实现对特定互联网资源的动态监测和实施抓取。第二,数据质量智能检查。通过制定规则,以机器自动检测为主,实现对数据完整性等质量指标进行核查。第三,多源数据的智能集成。企业底层资源的大数据尚未实现统一的组织和管理,海量的数据资源增加了用户获取细粒度知识的难度,不同的数据标准也导致了“数据孤岛”的存在,严重影响了企业大数据资源价值的有效过滤、发现和挖掘。通过制定统一的数据模型和元数据,规范和约束多源数据的数据结构,并采用数据结构自适应技术,实现多源异构数据的数据汇聚和集成,从而为情报服务的开展提供底层数据的智能化支撑。

3.2 基于领域技术研判的平台功能

传统基于人工对领域技术进行研判的工作,已不能适应大数据环境下的情报服务需求。第一,海量数据及信息处理已经超出人工处理的范畴,且人工对领域技术进行研判,不仅费时费力,而且最终结果不便于溯源和管理;第二,由于缺乏客观数据的支撑,专家存在主观认知偏差和知识盲区,由人工进行领域技术研判易产生偏差,不够全面;第三,面对不同技术类型的研判,由人工来处理加工,难以分析数据之间的语义关系,对数据分析的深度不够,容易造成结果的不准确;第四,线下开展的技术识别、跟踪工作,缺少统一平台进行展示和呈现,不便于共享。

面向不同业务和决策需求,采用文本挖掘、机器学习、知识发现等技术,基于海量数据对领域内的关键技术、核心技术、前沿技术、颠覆性技术等不同技术类型进行识别、评价和研判,全面分析领域内技术及其之间的关联,进一步发掘隐藏在软件数据背后的潜在和有价值信息^[28],实时动态展示不同技术主题变化趋势,把握领域技术发展走势,是企业技术创新的核心内容。

3.3 基于用户参与、精准智能的平台服务

企业情报用户的需求不仅呈现出多样化、多层面、个性化、专业化等特点,还需要以自治的、个性化服务的方式来满足需求,因此需要对传统“机械式”情报服务方式进行改革,在情报服务过程中强调用户全程参与的原则,从需求驱动到反馈设计都强调用户参与,逐步建立联动机制,从而满足用户的个性化情报需求。

置身于信息爆炸和信息冗余的环境,“千人一面”的情报服务方式难以适应多样化的情报需求,因此迫切需要建立大数据驱动的精准智能服务体系,研发多元情报产品,拓展情报服务模式,实现即时的个性化响应,将情报有效传达到用户,实现点对点的情报服务,满足情报用户的个性化服务诉求。此外,目前用户需要的是融合多种服务方式的情报服务,是一种具有“前瞻性”的情报服务需求,这就需要采用智能技术和智能工具,深度加工底层数据,识别和挖掘底层数据之间的语义关联,充分利用企业的静态和动态资源,为提供多样化的、深度的、精准的情报服务奠定基础。而且,现有的通用工具虽然也能够对一定数量的数据进行语义分析和加工,但面对专业领域内的数据,输出结果存在不能够解读、不够准确等弊端,因此需要提供专业化的情报服务,以满足用户的情报需求。

4 面向技术创新的企业智能情报服务平台研发

4.1 企业智能情报服务平台的信息生态圈

信息生态管理是智能情报服务实施的关键环节^[29]。企业智能情报服务平台的信息生态圈(见图1)从内到外由信息主体链、信息本体链和信息环境链组成。

(1) 信息主体链。信息主体链是由信息生产者、分解者、传递者、消费者4类信息主体形成的一个完整的闭环循环系统。信息生产者主要是指数据服务商,负责信息的采集和生产;信息分解者和传递者主要包括情报服务人员,需要具备情报分析、预测、服务专业技术,负责将信息进行组织、序化、加工,并且依据用户需求,筛选、集成和传递信息;信息消费者主要是指用户,对信息服务成果进行利用和评价。但这些信息主体的角色并不单一和固化,例如数据服务商还可以作为信息传递者担负信息组织、集成职责,为企业智能情报服务提

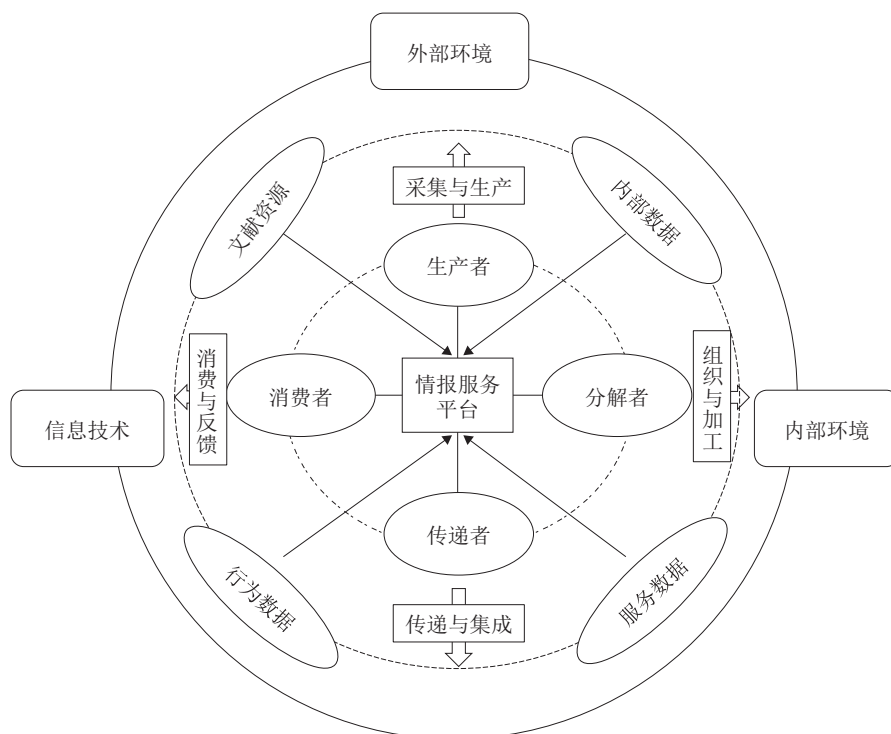


图1 企业智能情报服务平台的信息生态圈

供有力支撑。用户是智能情报服务的需求方,同时他们在平台中也充当专家角色,将自身知识和经验外化从而贡献知识成果,又可作为信息的生产者和传递者。

(2) 信息本体链。信息本体链涵盖的是各类数据资源,主要包括文献资源、内部数据、行为数据、服务数据等。文献资源优选高质量权威数据源,覆盖期刊、会议、学位论文、项目、报告等多种类型,为提供高水平的智能情报服务奠定丰富的数据基础;服务数据包括用户评价反馈、机构专家名单等;行为数据是指用户在平台上留下的行为数据,如用户访问日志、浏览时间、浏览次数等;对服务数据和行为数据的分析和利用,便于准确把握用户的情报需求,为个性化的情报服务提供依据;内部数据是由企业内部创造的数据,包括模型指标数据、技术点版本管理、技术点点评数据、领域词典等,作为智能情报服务正常开展的基础和保障。

(3) 信息环境链。企业内外部环境和信息技术构成该企业智能情报服务的信息环境链。内部环境是情报服务有序进行的基础,信息技术是系统的技术支撑,外部环境是情报服务合法运转的保障和约束。

内部环境是保障情报服务有序运转的基础设施、网络设备、服务平台、管理制度等,它们为信息资源的采集、加工处理、存储、数据分析、成果发布等一系列工作

提供基础保障和支撑,内部环境直接影响企业智能情报服务的服务质量、服务方式、服务流程和服务情境。

外部环境是对系统内部活动能够产生影响的要素,主要包括宏观的信息政策、信息法律、信息文化和伦理等,是情报交流、应用、分享、传播的合法性保障和约束,能够保障情报服务创新,保持健康良性的发展方向。

信息技术是保障信息流转及信息生态链的合理流动,用于开发、交流、管理、利用信息资源,使得信息传递得到扩大并延伸的技术方法、传播方法和传递途径。智能技术的迅猛发展为面向技术创新的企业智能情报服务提供了前所未有的机遇,提升了机构智能情报服务能力,拓展了机构智能情报服务的深度,改善了情报服务方式和质量^[30]。

4.2 企业智能情报服务平台的框架设计

本文基于信息生态理论构建企业智能情报服务平台的信息生态圈,并以此为基础,以用户需求为导向,研发面向技术创新的企业智能情报服务平台。搭建的企业智能情报服务平台框架(见图2)主要包括基础设施层、数据资源层和服务应用层3个层次,它们对应于智能情报服务平台信息生态圈中的信息环境、信息本体

和信息主体3个部分: ①基础设施层是智能服务的物质基础和技术支撑; ②数据资源层为创新服务提供文献资源等核心资源, 以及文本智能解析、自动分类、主题建模、图谱构建、关联挖掘等数据处理技术; ③服务应

用层是智能服务的顶层交互端口, 为用户提供大数据驱动的各类技术发现、技术监测、优势资源分析、个性化服务等情报服务。该平台既是为用户提供个性化服务的情报平台, 也是智能的用户自助服务平台。

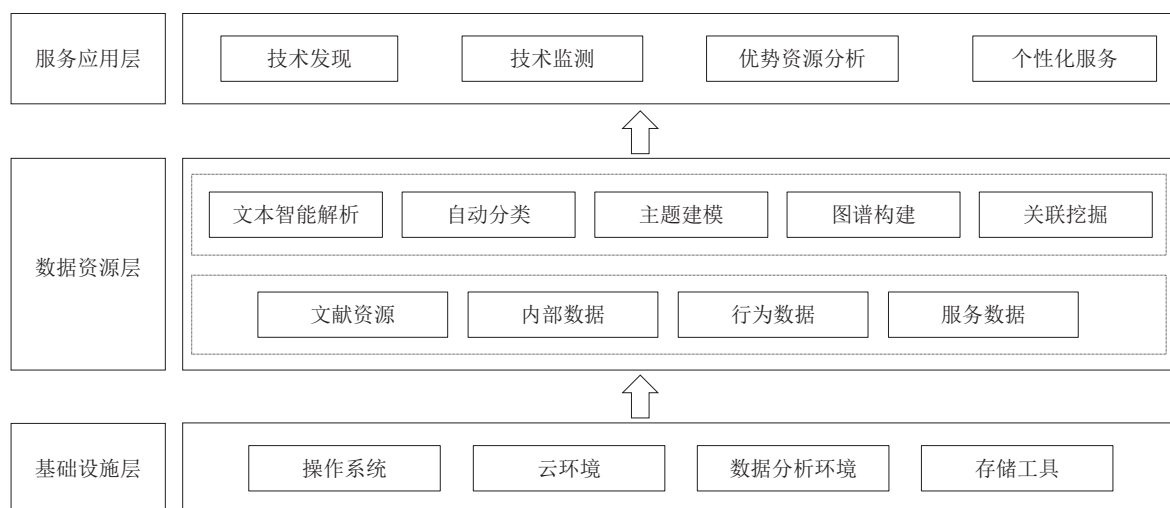


图2 企业智能情报服务平台框架

多层次的结构设计保证了平台在大数据环境下具有较强的适应性和可扩展性, 通过信息环境、信息本体、信息主体等组成部分的重新综合和分层, 可以揭示各要素在上下业务流程中的关联。一方面, 用户可以在满足需求的驱动下访问情报资源, 信息基础设施将获取、感知、识别和上传用户在访问资源时产生的服务数据、文献资源和行为数据等。情报服务人员可以利用数据挖掘、机器学习、人工神经网络等数据处理技术对上述数据进行管理、组织和分析, 实现不同的情报服务, 底层资源数据和服务数据的结合有助于用户作出高效准确的决策, 实现数据驱动决策功能。深入分析底层数据资源还可以帮助用户实现知识整合和知识发现, 提供发现隐性知识关联的服务功能。服务数据和用户行为数据有助于情报服务人员预测用户的个人需求, 并提供准确的知识推荐服务。

大数据驱动下面向技术创新的企业智能情报服务, 能够帮助用户实现从掌握知识、创新知识到获得智慧的过程。在此基础上, 用户在需求驱动下, 将更加积极地参与创建本地内容, 实现对高质量数据的沉淀。同时情报服务人员将进一步根据服务数据、业务数据、用户行为数据等, 提供精准、专业的知识推荐, 从而形成企业智能情报服务的循环发展生态链。

4.2.1 基础设施层

大数据的获取、存储、管理、组织、分析和应用依赖于稳定的基础设施支持, 基础设施层为企业智能情报服务的开展提供了一个必要的支撑环境, 实现了大数据驱动下的智能服务。设施包括一系列硬件感知的IT基础设施设备、存储设备、网络设备、监视器等。完善的数据感知识别系统具有纸质信息资源、数字信息资源等信息的深度感知、测量和捕获功能, 能够实时对大数据的全面感知、智能识别和实时上传。实现大数据驱动的企业智能情报服务, 还需要结合一系列智能技术来支持大数据处理, 从海量数据中实现聚类分析、统计分析、语义分析、预测分析、知识关联分析等功能。如大数据、云计算技术、区块链技术为数据资源存储、各类异构资源的整合提供了解决方案; 用户画像、个性化推荐和主动推送技术可以精准把握用户需求, 实现主动知识服务; 人工智能和机器学习技术能够实现知识萃取、技术挖掘、关系推理等功能, 为企业智能情报服务开展提供技术支撑和保障。

4.2.2 数据资源层

数据资源层负责底层数据的管理和分析, 基于底

层算力,综合运用文本智能解析、自动分类、主题建模、图谱构建、关联挖掘等智能分析方法,对文献资源、内部数据、行为数据、服务数据等数据资源进行深度加工、关联分析和重复利用,提升底层资源丰度和价值,探索用户行为模式,为面向技术创新的多个应用场景赋能。

4.2.3 服务应用层

服务应用层是整个系统的顶层交互端口,直接面向用户,通过多元化的产品形式,为用户提供智能情报服务。数据应用主要体现在大数据驱动的智能服务生态链的发展上,依赖于数据、情报服务人员、用户、平台、环境等生态要素的互动^[31]。平台为智能情报服务的开展提供了环境和数据,情报服务人员通过标准或规则,对海量数据进行序化、组织和整合,充分挖掘数据价值,作为提供智能情报服务的前提。用户在获得情报服务的基础上,可开展更高层次的知识创新活动,产生数据将进一步丰富底层数据,推动智能情报服务的开展和实施。

情报服务人员主要为企业用户提供4种技术创新情报服务:技术监测、技术发现、优势资源分析和个性化服务。技术监测服务主要是实时关注与领域技术相关的政策更新、行业发展、企业竞争、技术变革动态、周边动态等国内外资讯,并能够对重点主题提供检测预警。技术发现服务主要是能够对领域内的前沿技术、新兴技术、颠覆性技术、关键技术等不同技术类型进行识别,同时也能够对技术发展趋势、技术间关系进行研判和推断。优势资源分析是指筛选出领域内或区域内的优势技术、头部机构、TOP专家等资源,并进行可视化呈现。个性化服务是指用户可以定制感兴趣的主体,利用平台数据处理功能直接智能分析,全程零代码操作,能够快速获取分析结果。

4.3 企业智能情报服务平台的关键技术

4.3.1 多源数据整合技术

大数据环境下,技术相关数据涉及期刊、会议、专利、项目、资讯、科技报告、标准等多种文献类型,数据类型涉及字符串型、文本型、数值型等,数据结构庞杂。服务于不同企业的个性化情报服务系统间缺乏统一的

元数据标准,相同属性的元数据存在差异和不一致性,机械的数据整合会造成偏差,影响数据质量,因此明确各类数据的业务含义、数据元含义、数据项属性等内容,建立数据字段间的映射关系,并且根据业务含义,建立不同层次的主题域,明确不同层次主题域中的字段数量,从而建立统一的数据模型,便于数据流转和集成。

多源数据整合遵循如下步骤。首先,按照资源类型创建资源类型模型,并赋予唯一的资源类型ID。其次,针对每个资源类型ID,采用动态数据存储技术,自动感知和识别录入数据的字段名称、数据类型、长度范围。最后,对已入库的数据进行自动巡检,重点关注同一资源类型ID内,数据字段是否重复;制定去重规则,检查同一资源类型ID内,数据是否冗余,如存在数据重复,则删除重复数据。

4.3.2 主题聚类技术

文本主题聚类可以协助发现文本中蕴含的主题,是进行领域技术挖掘和识别的基础,常用的主题聚类方法有潜在语义索引(Latent Semantic Indexing, LSI)、隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)等主题模型。本平台以LDA主题模型为基础,在文本预处理、主题模型构建等方面进行改良,不仅提高了计算效率,而且操作简单易行,分析结果的准确性大大提升。

文本预处理阶段,建立句子级的句法切分框架,通过建立多个基于语法规则的上下文无关文法(Context Free Grammars, CFG),完成对英文文本分词操作。目前对英文文本语句拆分经常使用的是依据空格,或者使用CFG框架进行切分,这样容易造成短语的分裂,从而降低单词的语义内涵,而本平台使用的根据语句特点建立的语法规则CFG,不仅能使得单词的语义得到最大程度的保留,而且简单便于操作,占用内存小,容易实现工程化推广应用。

主题模型阶段,主题数、迭代次数、文档-主题先验参数 α 和主题-单词先验参数 β 的设置自动化,免除凭借人工经验的赋值,大大提升主题识别的准确性,降低了操作难度,为后续不同技术类型的识别和挖掘奠定坚实客观基础。主题数取困惑度曲线的峰值;迭代次数与主题数成正比,迭代次数 I 与主题数量 K 存在以下关系:当 $K \leq 10$, $I = 3000$;当 $K > 10$, $I = 3000 + (K - 10) \times 300$ 。

先验参数 α 和先验参数 β 的设置与迭代算法相关: 如果迭代算法采用的是EM方法, 参数 α 值为 $(50/k)+1$; 参数 β 值=1.1; 如果迭代算法采用的是贝叶斯方法, 参数 α 值取 $(1.0/k)$, 参数 β 取值为 $(1.0/k)$ 。

4.3.3 关系发现技术

通过对技术间关系的发现, 增加对技术信息的深度语义揭示, 是进行技术情报利用的关键环节。投入产出的APL (Average Propagation Length) 模型方法认为, 利用不同技术领域之间的引用关系, 可测量某项技术对技术网络的影响, 众多学者已将APL模型应用于突出多个科技领域的核心知识间的联系, 本平台在姜照华等^[32]发表的用于热点和技术突破机会的APL模型的基础上, 将APL模型和基于注意力机制的双向长短时记忆 (Bi-directional Long Short-Term Memory, BiLSTM) 关系抽取算法相结合, 用于技术关系发现。首先采用Attention+BiLSTM算法对技术点间的关系进行识别, 然后根据APL模型计算两两技术点间关系大小, 最后对关系大小设定阈值, 筛选出领域技术突破机会。

4.4 企业智能情报服务平台的核心功能

作为以识别和追踪技术发展前沿及其走势为目的的

企业级科技发展战略决策支持工具, 面向技术创新的企业智能情报分析平台主要实现以下5个方面的功能。

4.4.1 领域技术识别与评估

面对海量的文献资源, 传统的仅通过专家判别的人工技术点识别方法已经不能满足当前机构对技术点精准、快速的识别需求, 而机器学习和文本挖掘的大数据分析方法为这种需求提供了有效的解决方案。平台着重分析前沿技术、新兴技术、关键技术等不同技术类型特征, 研发多样的智能算法和构建模型群, 从中筛选最佳实现算法和评估模型指标, 达到识别和评估领域技术的目的。利用本平台对炼钢领域相关英文文献进行技术扫描, 通过系统内嵌的前沿技术评估模型, 选取技术创新性、技术成熟度、技术影响力3个指标, 得到炼钢领域技术识别及评估结果 (见表1)。各指标计算采用功效系数法, 值域设置区间为 (0~100), 模型总分的计算规则为 (技术创新性指标得分+技术成熟度指标得分+技术影响力指标得分)/3。本平台不仅直观呈现了直接还原铁技术、焦炉煤气喷吹、TRT发电技术、水淬高炉渣技术、转炉炼钢终点控制等技术在技术创新性、技术成熟度和技术影响力上的分别得分及模型总分, 而且还能够获取综合维度和每个维度的TOP技术, 从而协助研判技术的前沿性。

表1 炼钢领域前沿技术识别及评估结果

名 称	主题词	技术创新性得分	技术成熟度得分	技术影响力得分	模型总分
直接还原炼铁技术	direct reduction, dri, oxide, suspension rotor, hdri	64	95	68	76
焦炉煤气喷吹	coke oven gas, hydrogen, cog reduced	39	67	69	58
TRT发电技术	trt, top gas, ore powder, turbine unit	32	71	79	61
水淬高炉渣技术	water, granulated, blast furnace slag, water-quenched	57	79	75	70
转炉炼钢终点控制	endpoint, converter slag, prediction, endpoint, carbon content	78	89	78	82
复合吹炼技术	combined blowing, oxygen, top-bottom, converter	80	90	85	85
竖炉工艺	shaft furnace, gas, coal, hydrogen	59	76	83	73
熔融还原技术	smelting reduction, core, blast, corex	70	87	91	83
溅渣护炉技术	slag, lance, oxygen, splash	49	54	67	57
煤气干法除尘	dry dust, removal, gas, blast furnace	65	53	74	64

4.4.2 领域技术关系发现

为企业技术创新提供有价值的情报, 不仅需要提供领域技术的名称, 还应尽可能提供技术的研发路径。

为了提供全面、精准的决策支撑情报, 本平台对领域技术间的关系进行识别和挖掘, 通过计算技术间的相似度, 不仅能够了解技术间的关联程度, 而且还能够为技术突破机会的发现提供预见和指导。以炼钢领域为例,

本平台可提供炼钢领域技术间关系图(见图3),节点大小表示对应领域的热度,连线粗细表示关系大小。图3中设置阈值为0.6,炼钢领域热度TOP5依次为复合吹炼技术、直接还原炼铁技术、氢冶炼、转炉双渣工艺、

焦炉煤气喷吹。领域间技术突破机会TOP5为复合吹炼技术与转炉双渣工艺、铁水预处理与直接还原炼铁技术、铁水预处理与氢冶炼、转炉与溅渣护炉技术、复合吹炼技术与干熄焦技术。

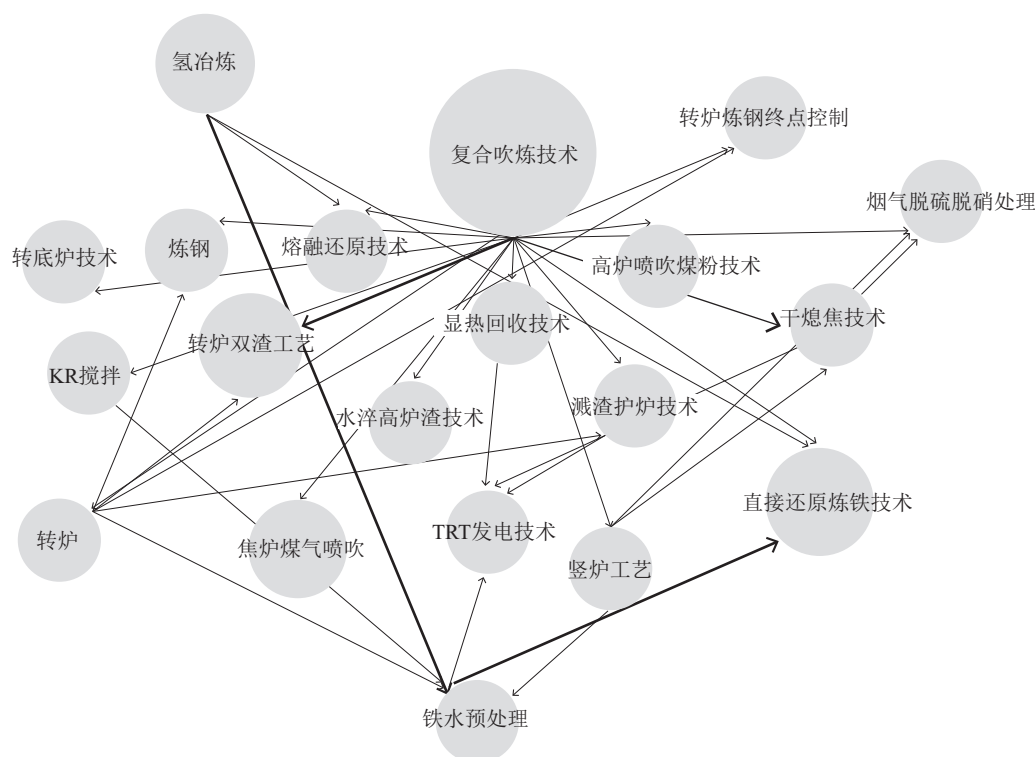


图3 炼钢领域技术间关系图

4.4.3 重点企业研发布局挖掘与追踪

企业进行技术创新不仅要了解行业和技术的前沿,而且还要对典型企业的技术动态和布局进行识别和追踪,从而能够知己知彼,有助于明确研发方向和战略部署。以典型企业为核心,通过网络爬虫技术、科学文献计量、自动分类、数据可视化分析、知识图谱分析等技术实现企业技术布局、技术关系的展示,并且能够对技术相关的资讯、政策、专利、期刊、项目等信息进行持续监测和一站式集成,分析呈现技术热点和技术发展趋势,全方位提供企业技术情报画像。通过本平台可对钢铁领域头部企业韩国浦项制铁公司(POSCO)的技术布局、发文趋势、技术动态等信息一目了然地展示和呈现。

4.4.4 用户自助智能服务

将情报分析中常用的工具和模型,如语义分析、自

动分类、主题聚类情报工具虚拟化,可以帮助有需求的用户在平台上自助实现对特定主题的个性化、按需分析。

用户可以将自己感兴趣的主体,通过确定主题词,采用直接搜索或自助式上传的方式,获取目标数据。用户可以在Web端通过零代码,实现对底层资源自动分类、主题聚类、实体识别、关系抽取、关键词抽取等数据分析,从而快速、准确、客观地了解到关于特定主题的技术体系,对不同技术类型的评估,以及在该主题领域中的优势资源,实现对该领域技术全方位的洞察。

4.4.5 “一站式”科技情报决策平台

“一站式”科技情报决策平台可从多个方面进行理解。第一,对于单个用户来说,用户通过该平台可以完成从数据采集、数据预处理、数据处理与加工、数据动态追踪、分析报告下载等全链条情报服务;第二,对于企业来说,不同用户可以登录该平台,协同共享特定

领域的技术情报分析成果,从而确保技术情报信息来源的统一性;第三,通过该平台,可以实现对不同领域主题技术体系的管理,以及洞察技术体系结构演变,便于企业的知识管理。该平台可以打通上下游,促进创新主体跨领域、跨行业协同创新,突破资源能力限制,提高创新绩效。

5 结语

技术情报是支撑企业创新发展的重要资源,而面向技术创新的企业智能情报分析平台的搭建对于企业高效开展技术情报有重要价值。本文设计并成功开发了面向技术创新的企业智能情报分析平台,以此作为科技发展态势及前沿技术趋势分析的工具,为科技发展战略和科技政策的制定提供决策支持。该平台通过高效获取和融合不同来源、不同格式的信息,运用文本聚类技术、自动分类技术,深入挖掘领域技术主题,以及技术主题领域的优势资源,结合国家、时间等维度的变化,重点分析每个技术点的研究热点、研究趋势,为分析科技发展的宏观走势,以及对科技、经济和社会产生影响的科技发展前沿趋势提供支撑。当然,随着科技发展和技术进步,根据用户需求,技术情报平台功能在实际使用过程中还需要不断完善与拓展,智能情报服务仍需不断升级。

参考文献

- [1] MASON R M. Introduction: The Challenge of Information Service Management [M] //Information Services: Economics, Management, and Technology. London: Routledge, 2019: 1-5.
- [2] 张凤海,徐丽娜,侯铁珊.关于技术创新概念界定的探讨[J].生产力研究,2010(3):26-27.
- [3] 钱旭潮,范苗苗.企业技术创新成长阶段的内涵、特征与条件[J].管理现代化,2019,39(3):52-54.
- [4] 巫英.基于德尔菲调查法的能源领域技术预见研究[J].武汉理工大学学报(信息与管理工程版),2017,39(3):338-341.
- [5] 安达,李梦男,许守任,等.中国工程科技信息与电子领域2035技术预见研究[J].中国工程科学,2017,19(1):50-56.
- [6] “中国工程科技2035发展战略研究”农业领域课题组.中国工程科技农业领域2035技术预见研究[J].中国工程科学,2017,19(1):87-95.
- [7] 许学国,桂美增.基于机器学习的新能源汽车核心技术识别及布局研究[J].科技管理研究,2021,41(9):96-106.
- [8] 周源,刘宇飞,薛澜.一种基于机器学习的新兴技术识别方法:以机器人技术为例[J].情报学报,2018,37(9):939-955.
- [9] 王秀红,高敏.基于BERT-LDA的关键技术识别方法及其实证研究——以农业机器人为例[J].图书情报工作,2021,65(22):114-125.
- [10] 周海炜,吴成凤.基于专利SAO结构和多指标评价的新兴技术识别研究——以手机芯片领域为例[J].情报杂志,2022,41(2):48,86-94.
- [11] 李勇,欧志梅,黄格,等.基于专利挖掘的数字文旅关键技术识别和趋势分析[J/OL].图书馆论坛:1-12[2022-05-05].<http://kns.cnki.net/kcms/detail/44.1306.G2.20220309.1005.002.html>.
- [12] 王燕鹏,韩涛,陈芳.融合文献知识聚类 and 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.
- [13] 鲁啸,王健美,卢絮.企业技术情报服务平台原型设计——以某风电企业技术情报平台建设为例[J].竞争情报,2021,17(1):29-35.
- [14] 刘民.电力企业竞争情报服务平台的研究与实现[J].计算机时代,2019(3):98-100,104.
- [15] 李子林,李雪山,郭肖肖,等.我国轨道交通行业科技情报服务创新研究[J].数字图书馆论坛,2021(9):43-48.
- [16] 姜岩,张赏,黄鲁成,等.基于专利分析的替代性技术识别研究[J].情报杂志,2014,33(9):27-32.
- [17] 吴菲菲,杨梓,黄鲁成.基于创新性和学科交叉性的研究前沿探测模型——以智能材料领域研究前沿探测为例[J].科学学研究,2015,33(1):11-20.
- [18] 黄鲁成,成雨,吴菲菲,等.关于颠覆性技术识别框架的探索[J].科学学研究,2015,33(5):654-664.
- [19] 罗建,蔡丽君,史敏.基于专利的两阶段新兴技术识别研究——以图像识别技术为例[J].情报科学,2019,37(12):57-62.
- [20] 李欣,王静静,杨梓,等.基于SAO结构语义分析的新兴技术识别研究[J].情报杂志,2016,35(3):80-84.
- [21] 宋凯,朱彦君.专利前沿技术主题识别及趋势预测方法——以人工智能领域为例[J].情报杂志,2021,40(1):33-38.
- [22] 黄鲁成,刘春文,吴菲菲,等.基于NPCIA的核心技术识别模型及应用研究[J].科学学研究,2020,38(11):1998-2007.
- [23] 王明强,李国正,陈欣然,等.中医药智能服务技术预见与关键技术识别研究[J].医学信息学杂志,2022,43(4):64-70.
- [24] 崔怡雯,赵筱媛,苏成,等.面向颠覆性创新的技术监测分类体系研究[J].情报学报,2021,40(12):1288-1293.
- [25] 臧冀原,刘宇飞,王柏村,等.面向2035的智能制造技术预见和路线图研究[J].机械工程学报,2022,58(4):285-308.

- [26] 崔蓓, 王磊. 我国生物医药关键技术预见研究 [J]. 医学研究杂志, 2022, 51 (3): 5-10.
- [27] 栾春娟, 李宇婧. 全球能源产业的技术演进与技术预见 [J]. 技术与创新管理, 2022, 43 (1): 10-20.
- [28] CUI Z, YAN C. Deep integration of health information service system and data mining analysis technology [J]. Applied Mathematics and Nonlinear Sciences, 2020, 5 (2): 443-452.
- [29] SANG W Z, HE P, YU H J. Research on Knowledge Governance Mechanism of B2C E-commerce Supply Chain from the Perspective of Information Ecology [C] //2019 4th International Conference on Social Sciences and Economic Development (ICSSSED 2019). Atlantis Press, 2019: 72-79.
- [30] LIU C, DOU T, ZHOU H, et al. Library service innovation based on new information technology: taking the interactive experience space “Tsinghua Impression” as an Example [J]. International Journal of Library and Information Services, 2021, 10 (1): 71-81.
- [31] QU J. Research on mobile learning in a teaching information service system based on a big data driven environment [J]. Education and Information Technologies, 2021, 26 (5): 6183-6201.
- [32] 姜照华, 宋华东, 王静, 等. 技术网络中新兴技术的突破机会的 APL识别方法 [J]. 科学学与科学技术管理, 2020, 41 (12): 71-82.

作者简介

冯向梅, 女, 1984年生, 博士, 工程师, 研究方向: 企业竞争情报服务、知识服务产品研发, E-mail: fengxm@cncic.cn。

顾方, 男, 1982年生, 高级工程师, 研究方向: 企业竞争情报服务。

鲁瑛, 女, 1977年生, 高级工程师, 研究方向: 信息资源管理。

周群芳, 女, 1973年生, 硕士, 副研究员, 研究方向: 企业竞争情报、专利分析与预警。

王俊松, 男, 1980年生, 高级工程师, 研究方向: 竞争情报系统研发。

肖甲宏, 男, 1988年生, 硕士, 工程师, 研究方向: 企业竞争情报服务。

Construction of an Intelligent Information Service Platform for Enterprises Based on Technological Innovation

FENG XiangMei¹ GU Fang¹ LU Ying¹ ZHOU QunFang² WANG JunSong¹ XIAO JiaHong¹

(1. Information Service Department, China National Chemical Information Center Co.,Ltd, Beijing 100029, P. R. China; 2. Central Research Institute, China BaoWu Steel Group Corporation Limited, Shanghai 200126, P. R. China)

Abstract: This paper designs and implements a “one-station” enterprise intelligent information analysis platform for technological innovation with fast response, in-depth analysis, integrated functions, flexibility, and visualization presentation. The platform has many functions, including multi-sources data fusion, text clustering, relationship discovery, other intelligent technologies to achieve intelligent, efficient and accurate technology information acquisition, technology identification and evaluation of different subject areas, relationship discovery of different technology areas, key enterprise R&D layout mining, and user self-service intelligent services. It aims to provide enterprise users with efficient strategic decision support services, in terms of knowledge structure analysis, development trend and advantageous resources of domain technologies, and various information visualization. It also can provide effective and convenient research methods and analysis tools for enterprise technology intelligence research business, and improve intelligence service capability and efficiency.

Keywords: Multi-Sources Data; Relationship Discovery; Frontier Technology; Technology Foresight; Information Ecology

(收稿日期: 2022-04-12)