

专利信息分析系统分析与研究

1. 厦门大学萨本栋微米纳米科学技术研究院 厦门 361005;
2. 厦门大学信息科学与技术学院 厦门 361005;
3. 同方知网数字出版股份有限公司 北京 100083

康婧¹ 谢怡² 宋佳颖¹ 赵正青³ 张庆国³ 康怀志¹

摘要 专利是创新成果的重要载体,对专利资源的有效利用有助于打破信息壁垒,准确评估成果的创新性和价值,推动科技成果向生产力转化。专利信息分析系统可以提高工作效率和创新能力,完成专利数据收集、专利信息检索、专利技术挖掘和分析等工作。本文对专利信息分析系统的功能和分析算法进行了分类介绍,并结合典型系统进行展示和比较,同时提出专利信息分析系统面临的问题与挑战。

关键词: 专利信息分析系统, 科技创新, 专利检索, 专利分析, 信息挖掘

中图分类号: G35

开放科学(资源服务)标识码(OSID)



A Review of Patent Information Systems

1. Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, Xiamen 361005, China;
2. School of Information Science and Engineering, Xiamen University, Xiamen 361005, China;
3. Tongfang Knowledge Network Digital Publishing Technology Limited by Share Ltd, Beijing 100038, China

KANG Jing¹ XIE Yi² SONG JiaYing¹ ZHAO ZhengQing³ ZHANG QingGuo³ KANG HuaiZhi¹

Abstract As important carriers of innovation, patents are usually used to break information barriers, evaluate scientific achievements and drive the transforming from technical innovation to productivity as

基金项目: 本研究得到科技部创新方法工作专项(No. 2015IM020500)的资助。

作者简介: 康婧(1987-), 硕士, 助理工程师, 研究方向: 专利信息分析系统; 谢怡(1979-), 博士, 助理教授, 研究方向: 网络与数据库, 计算机应用; 宋佳颖(1992-), 硕士, 研究实习员, 研究方向: 环境工程; 赵正青(1976-), 本科, 高级工程师, 研究方向: 知识管理; 张庆国(1976-), 硕士, 高级信息系统项目管理师, 研究方向: 文本挖掘、中文智能信息处理; 康怀志(1975-), 通讯作者, 博士, 教授级高级工程师, 研究方向: 拉曼光谱技术研究与应用、生物纳米技术研究以及微小型分析检测仪器技术研发等, Email: kang@xmu.edu.cn。

important carriers of innovations. Patent information can improve the innovation ability by increasing the working efficiency of searching and analyzing patent information. This review introduces the major functions and algorithms of patent information systems, and also presented and compared the representative systems. Moreover, the issues and challenges for patent information systems are discussed in this study.

Keywords: Patent information system, technical innovation, patent search, patent analysis, data mining

1 前言

近年来,国家高度重视科技创新,“十三五”国家科技创新规划进一步明确了深入实施创新驱动发展的战略^[1]。在知识产权战略背景下,专利知识产权成为科技创新的重要载体和驱动力,专利资源的重要作用贯穿了科技创新的全过程^[2]。研究人员通过获取国内外和其他机构的相关专利资源,打破信息壁垒,判断目前工作的原创性,避免研究资源的浪费;同时通过学习现有的专利技术并进行吸收和再创新,最终实现自主创新。专利代理人和专利审查员在判断科研成果的创新性是否足以申请专利授权时,也需要对比其他专利,而这些工作都建立在专利资源的有效管理和分析基础上。对专利资源的有效利用还能加速研究成果转化为专利权的过程,推动科技创新成果向生产力的转化。

随着全球专利申请量的激增,专利资源越来越丰富,撰写和审查压力也越来越大。为了有效利用专利资源开展研发活动,缓解专利撰写和审查压力,人们借助专利信息系统提高工作效率。研发人员通过专利信息系统准确检索专利信息,全面分析和评估相关核心技术,例如分析专利战略布局、挖掘具有研究价值的技术空白、评估对现有专利二次创新的价值等。

专利代理人和专利审查员,主要关注申请专利是否和其他专利相似,因此需要在数据及时更新的专利信息系统中进行全面准确的检索。

因此,专利信息系统主要包括三大功能模块:专利检索功能,专利分析功能和专利评估功能。较成熟的专利信息系统均具备这些功能模块,例如 Thomson Innovation^[3], Incopat 科技创新情报平台^[4], Patsnap^[5], 中国知网专利系统(知网系统)^[6]和 Innojoy 专利搜索引擎^[7]等。本文对专利信息系统的重要功能和算法进行了详细介绍,并以上述专利信息系统为典型代表,描述具体功能,帮助用户更好地了解系统的设计和特点。同时讨论专利信息系统面临的问题和挑战,并提出解决思路。

本文主体内容的结构如下:第二部分详细介绍专利信息系统的三大功能模块及重要功能;第三部分对比典型专利信息系统的功能和特点;第四部分是讨论专利信息系统面临的问题及挑战;第五部分进行总结。

2 专利信息系统的主要功能

2.1 专利检索功能模块

专利检索功能模块的作用是对专利资源的定位和提取,主要功能包括简单检索、表格检索、逻辑检索、IPC 分类号检索、语义检索、引证

检索和法律状态检索等；辅助功能包括二次检索和去重检索。

简单检索是根据输入的关键词对专利的著录项和全文进行全面检索以定位和提取专利。表格检索是指将专利关键词按著录项进行检索，每一类关键词只对应某个著录项。逻辑检索则是对应各著录项的关键词采用逻辑表达式进行精准的检索。逻辑检索兼具表格检索和简单检索的优点，可通过构建表达式获得指向清晰的检索结果，灵活性强于表格检索，准确性优于简单检索。正确有效地运用这三种基础功能可

以实现专利信息检索的全面性和准确性，即查全率和查准率^[8]。

其他检索功能可提升用户效率，提高查全率和查准率。语义检索是指输入所需专利技术信息相关的词、语句、段落和篇章以检索相关专利。引证检索是指通过专利的引证和被引证关系检索专利。IPC分类号检索是通过专利的IPC分类号检索专利。法律状态检索是通过不同法律状态对专利进行筛选。如表1所示，Incopat和Patsnap系统的检索功能更为全面。

表1 专利信息系统的检索功能对比

系统名称	Incopat	Patsnap	Thomson Innovation	Innojoy	知网系统
简单检索	√	√	√	√	√
表格检索	√	√	√	√	√
逻辑检索	√	√	√	√	√
语义检索	√	√	×	×	×
引证检索	√	×	×	×	×
IPC分类号检索	√	√	×	√	×
法律状态检索	√	√	×	√	×
二次检索	√	√	√	√	√
去重检索	数据库已去重	√	√	√	数据库已去重
其他检索功能	扩展检索	图像检索、化学检索、扩展检索	×	同义词检索	×

由于配套数据源加工深度不同，各项检索功能所支持的字段和格式存在一定差异。对于基本的简单检索、表格检索和逻辑检索功能，常用的可检索字段包括发明人、专利权人、国家代码、公开（公告）号、公开（公告）日、授权号、授权日、优先权号、优先权日、IPC分类号、申请号、申请日、主分类号、代理人 and 任意词检索。这些检索字段是查全率和查准率的基本保障，属于必备字段。而IPC检索功

能只能检索IPC分类号；法律状态检索功能只能检索法律状态；引证检索功能只能检索专利引证和被引证情况。此外，Innojoy、Patsnap和Incopat还可检索专利代理机构、地址、最新法律状态、最新法律状态公告日、国际申请、国际公布、美国专利分类号和联合专利分类号等字段。Incopat的可检索字段最多（超过230字段），而Patsnap在分类号上增加了日本FI分类号^[9]和F-term分类号^[9]，同时将申请人和发

明人做了标准化，避免因同名误判而降低专利的查全率和查准率。

总之，数据源的有效加工和检索字段的灵活设计，可以提高查全率和查准率，可为后续专利的正确分析和评估提供保障，减少重复开发，提高二次创新的能力。因此，设计专利信息分析系统时，应根据用户需求对字段进行选择 and 配置，并对数据源进行加工。

2.2 专利分析功能模块

专利分析功能模块采用数学、统计学、社会网络关系学和机器学习等方法，深度分析专利技术相关信息（如技术竞争情报和领域成熟度等），

以此为基础获得宏观决策建议^[10]。统计类分析对专利分析指标进行数量统计，主要包括趋势分析、申请人分析、地域分析、法律及运营分析、技术分析等功能^[11-12]。引证类分析揭示专利之间的引用和被引用关系，主要包括专利互引网络分析^[12-14]、引证和被引证分析等功能^[14]。聚类分析^[15]通过聚类手段将专利文本映射到二维平面上，越相似的专利文本距离越接近，主要功能包括技术主题聚类^[16]、专利地图^[17]和关键词共现分析等^[18]。社会网络类分析借助社会网络分析手段研究专利相互关系，例如作者合作网络分析^[13]。表2对常见的专利信息系统的分析功能进行对比，分析结果由图或表格展现。

表2 专利信息系统的分析功能对比

功能类别	具体功能	Incopat	Patsnap	Thomson Innovation	大为 innojoy	知网系统
统计分析 (基本)	趋势分析功能	√	√	√	√	√
	技术总体分布	×	×	√	×	×
	专利组合分析	×	×	√	×	×
	授权概率分析	×	×	×	√	×
	发明人分析	√	√	√	√	×
	申请人分析	√	√	√	√	√
	分类分析	√	√	√	√	×
	代理机构分析	×	√	√	×	×
	地域分析	√	√	√	√	√
	法律及运营分析功能	√	√	√	√	×
引证分析	技术分析	√	√	√	√	√
	引证和被引证分析	√	√	√	√	×
	专利互引网络分析	×	×	×	×	√
聚类分析	技术主题聚类	√	√	√	√	×
	专利地图	√	√	√	√	×
	关键词共现分析	×	×	×	×	√
社会网络分析	技术功效矩阵	√	√	×	×	×
	作者合作网络	×	×	×	×	√

2.2.1 统计分析

统计分析功能是专利信息系统的基本功能，以 Incopat 系统的统计分析功能为例。图 1 显示的分析结果对应一个图和一张表，用户既能从图

中获得可视化信息（图的类型可定制），又可从表中获得准确数据（图表对照）。其中，申请人分析功能对常用信息项目进行统计，例如申请人排名、申请人类型构成、申请人申请趋势、申请

人公开趋势和申请人技术构成。法律及运营分析功能根据不同专利当前法律状态进行分析。代理

机构分析功能则针对代理机构 / 代理人排名、代理机构 / 代理人类型构成进行分析。

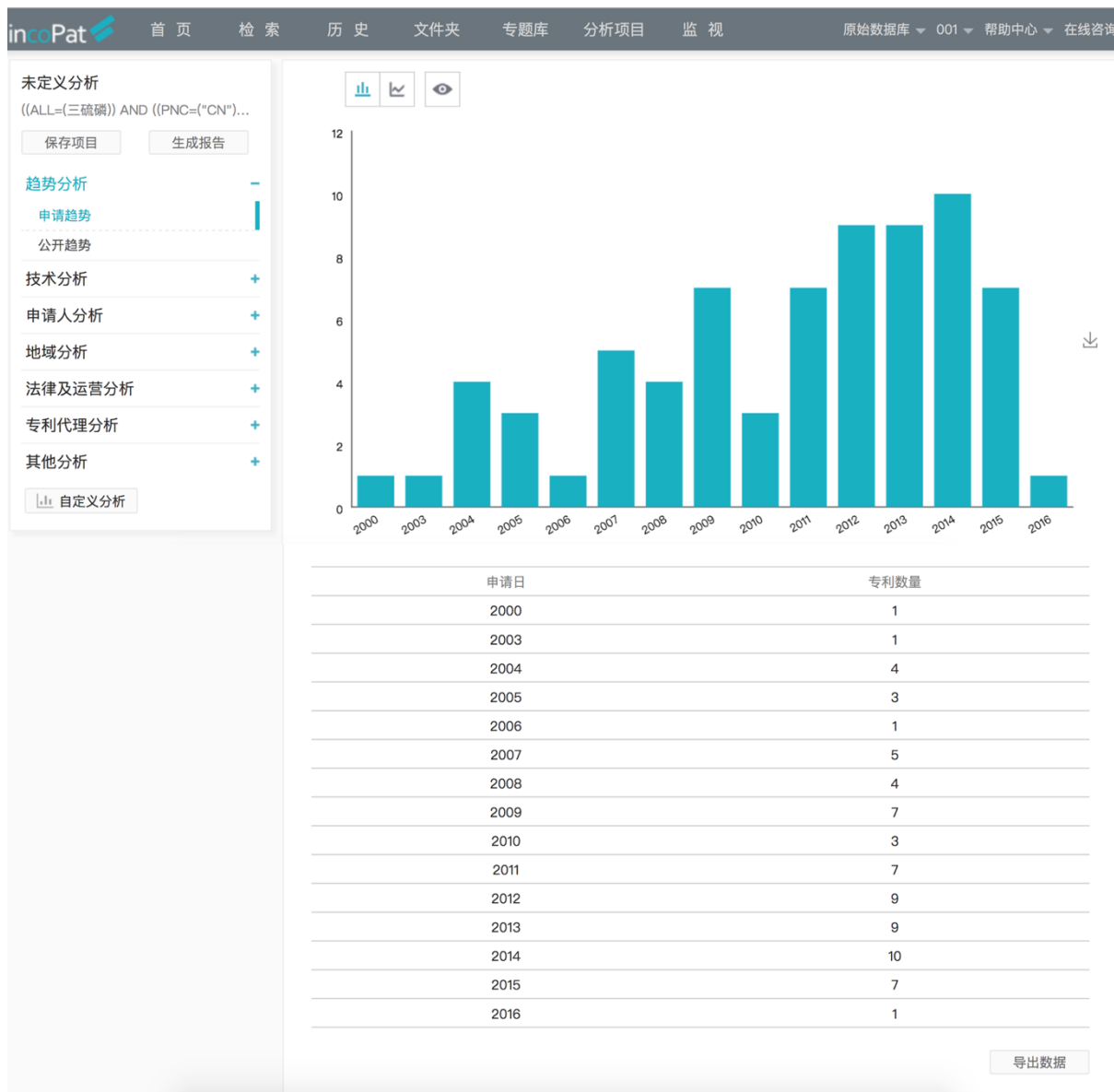


图1 IncoPat: 基本统计功能界面

此外, IncoPat 和 Patsnap 还提供了统计分析的扩展功能, 主要包括:

(1) 在申请人 / 发明人分析方面, 开展发明人申请趋势、发明人技术分布、申请人专利类型、申请人专利价值、申请 / 授权时间等分析。

(2) 在法律状态方面, 开展专利有效性、转让趋势、转让人排名、受让人排名、转让技术构成、许可趋势、许可人排名、被许可人排名、许可技术构成、诉讼当事人排名、复审无效请求人排名、无效宣告专利权人排名等信息分析。

(3) 在分类分析方面,开展 IPC 分类排名、IPC 分类申请趋势、UPC 分类排名、UPC 分类申请趋势、LOC 分类排名、LOC 分类申请趋势、CPC 分类排名、申请人 IPC 技术对比、发明人 IPC 技术对比、申请人 CPC 技术对比分析。

(4) 在地域分析方面,开展省排名、申请人国家申请趋势、申请人国家公开趋势、各省申请趋势、各省专利类型分析。

(5) 在代理机构分析方面,开展代理机构排名、代理人排名、代理机构申请趋势、代理人申请趋势、代理机构专利类型、代理人专利类型、代理机构 IPC 技术分布、代理人 IPC 技术分布分析。

2.2.2 引证分析

引证分析用树状图或网络状图来表示。其中,专利互引分析功能是指两个或多个专利因同时引用了同一个专利所形成的网络图。因为核心专利一般处在互引网络的中心,所以可以通过专利在网络图中的位置判断专利的核心程度。引证和被引证分析结果由树状或扇面状图来表示。通过统计某个专利的引证和被引证的专利数目,可以获得核心专利及其核心技术等技术竞争情报信息,更好地吸收核心专利技术进行二次创新,并通过专利布局规避对手专利。通过分析专利的引证关系还可以追溯某个领域专利技术的起源,了解领域专利的发展趋势,从而更好地把握领域技术的脉络,有效地帮助研发和专利宏观战略规划。目前,Incopat、Patsnap、Thomson Innovation 和 Innojoy 系统均配备了专利引证和被引证分析功能,知网系统也具有文献互引分析功能。

2.2.3 聚类分析

聚类分析结合数学和机器学习算法将专利文本映射到二维平面上,越相似的专利文本距离越接近,其结果一般用图表示。常见的聚类算法主要包括 k-means^[19] 和 BIRCH^[20] 等算法。k-means 算法将 N 个技术主题根据相似度做出初始赋值,然后将其分为 L 个分组进行聚类,并采用迭代的手段对聚类结果进行优化,使同一分组内的结果更接近而不同分组内的结果距离更远。BIRCH 算法是一种层次聚类方法,通过在聚类特征树中插入子聚类,然后根据阈值对聚类特征树进行重新构造,之后再行划分聚类分析。图 2 展示了技术主题聚类功能,可以通过改变技术主题聚类图的横纵轴指标获得不同聚类结果,其中 X 轴是不同的聚类主题(可自选关键词),Y 轴可以按申请/公告日、申请人、标准申请人、IPC、法律状态、地域、申请人国别、专利公开国进行分类。

聚类分析结果还有更直观的表现形式--专利地图。如图 3 所示,Patsnap 采用地图的绘制方法,通过对不同专利权人的同种技术方案进行技术主题聚类,将聚类结果投影到 3D 空间中:山峰部分是技术热点而山谷部分是技术空白点,柱状代表专利数目排名在前的若干专利权人(或发明人、分类号、国际专利公开号)所申请的各个专利。系统已具备文本挖掘功能,自动挖掘每个聚类中的关键字作为地图中每个专利领域的关键字(亦是领域技术的标识),同时支持选择细分领域后的进一步聚类分析。此外,地图上还对诉讼专利,高价值专利和授权专利做了标识,为后续规避纠纷和申请专利提供了技术竞争情报。知网系统通过聚类分析实现的

关键词共现分析功能，功能较为灵活。如图4所示，可以通过调整共现频次，节点过滤（共

现频次筛选），分析临近节点关系，年份选择和聚类数，选择获得关键词间的共现关系。



图2 Incopat：技术主题聚类功能界面

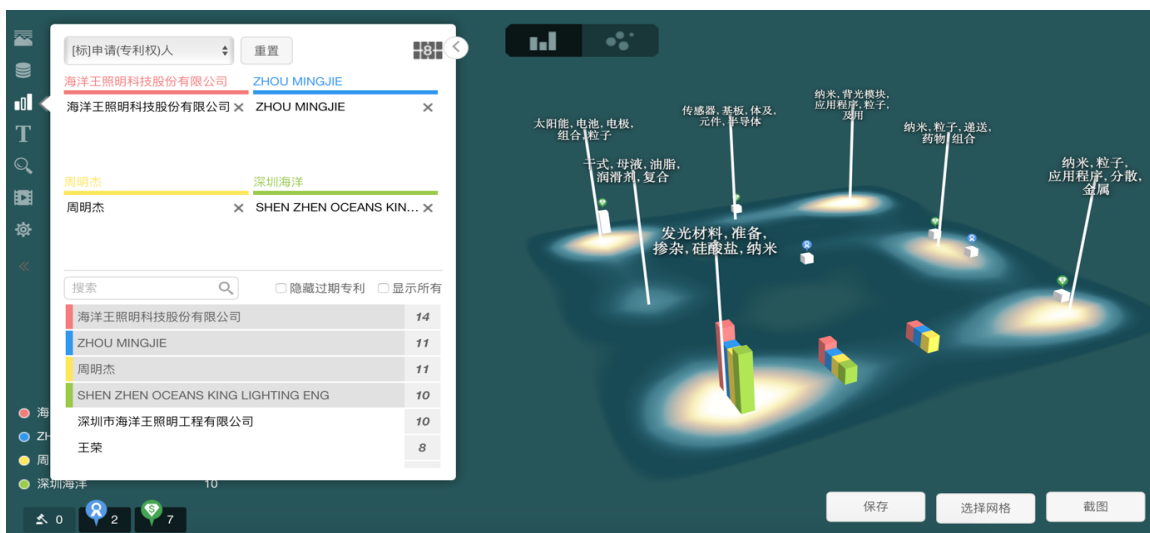


图3 PatSnap：专利3D地图界面

关键词共现网络

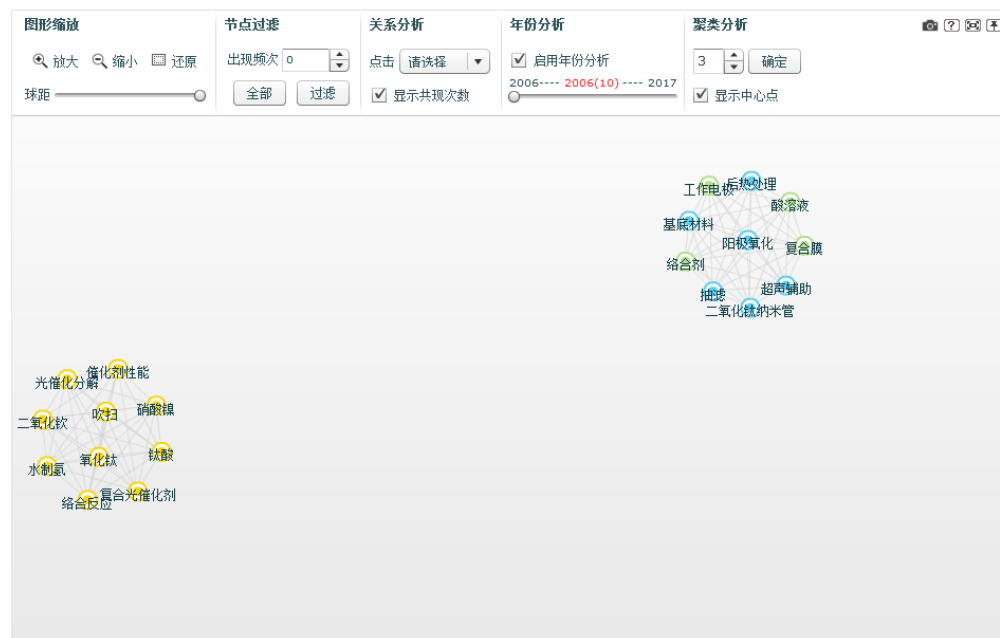


图4 知网系统：关键词共现分析功能界面

聚类分析还能将对应专利技术的功效关键词放入技术功效矩阵^[21]，从而发现研究的空白点。技术功效关键词可以通过建立模型、采用算法或通过其他自动手段抽取^[22]。人工抽取技术功效关键词更符合研发人员和专利工程师的实际需求，但缺点是工作量太大。因此，Patsnap 在关键词生成矩阵方面进行了很多自动化改进，而 Incopat 则设计自动实现了类似技术功效矩阵的功能。

2.2.4 社会网络分析

社会网络分析结果可用网络图直观表示。作者合作网络进行无标度网络分析，展现发明人地理空间分布网络分析、拓扑学结构和网络演进分析等结果^[13]。如图5所示，可选分析手

段包括图形缩放，节点过滤（共现频次筛选），关系分析，年份分析，并通过调整作者共现频次，年份选择和分析临近节点作者关系来多方面获得作者间的合作关系。

2.3 专利评估功能模块

专利评估功能模块的主要用途是对核心专利和核心技术的定位及评估，主要实现方式为利用多个评估指标综合后的结果对专利的重要性进行排序，然后筛选出排序在前的核心专利。核心专利的识别有利于对核心技术和核心创新点进行重点分析，是通过对单个专利评估分析寻找研发重点、空白点和创新点的有效参考。主要功能包括单参数评估功能和多参数评估功能，如表3所示。

作者合作网络分析



图5 知网系统：作者合作网络功能界面

表3 专利信息评估类功能对比

系统名称	Incopat	Patsnap	Thomson Innovation	Innojoy	知网系统
引用和被引用次数	有此参数但无对应单参数评估功能	有	有	有此参数但无对应单参数评估功能	有
同族专利个数		有	有		有
存活期		无	无		可计算
特色功能	合享价值度，结合多个评价指标通过设置的权重计算综合评分	无	引证树查找核心专利	DPI 大为价值度指数评分，结合多个单一评价指标通过设置的权重计算	无

主要的单参数包括：引用次数、被引用次数、同族专利个数和存活期。专利的引用和被引用次数是指专利引用其他专利或被其他专利引用的次数。同族专利^[23]是指由至少一个共同优先权联系的一组专利文献，称一个专利族（Patent family）。在同一专利族中每件专利文献被称作专利族成员（Patent family members），同一专利族中每件专利互为同族专利。专利存活期是指专利从授权到失效的时长。通常来说，引用和被引用次数越高，同族专利个数越多，存活期越长，则专利越核心。

多参数评估功能是指结合引用和被引用次

数，同族专利个数，存活期等主要参数以及权利要求数、说明书页数、发明人数等其他参数，通过专家赋值，算法赋值和用户自赋值的方式对这些参数赋予一定的权重，最后采用数学公式计算出专利的价值度。例如 Incopat 根据技术稳定性、技术先进性和保护范围三个部分的指标评分，可以得到专利的合享价值度的综合评分（如图6）。其中，技术稳定性的评价指标包括专利的授权状态，诉讼质押，复审无效等法律状态；技术先进性的评价指标包括其被引用次数，IPC 领域，发明人数和许可转让状态等；保护范围的评价指标包括权利要求和同族专利等。由于综合了多个有

效评价指标，并采用适宜的权重，对专利核心程度的评价较为可靠，现有的专利信息系统，也越

来越倾向采用多参数评估功能，以获得更可靠的专利价值度评估结果。



图6 Incopat的专利评估功能界面

3 专利信息系统的对比和经验借鉴

专利信息系统的基本功能大体相同，但功能全面性、灵活性和准确性各有不同。一般的，功能越全面的专利信息系统价格可能越高。因此，评价和选择一个专利信息系统，需要结合其性价比和用户需求来进行综合评判。根据对典型专利信息系统的功能对比分析，本文列出以下特点和经验，以供借鉴。

1. 各专利信息系统的数据库数据量和更新频率各不相同。数据来源主要是德温特数据库^[24]，有些系统（例如 Incopat 和 Patsnap）也从其他数据库获取数据，因此数据更新频率从日更新到双周更新不等。数据来源越广泛，需要整理和处理的数据就越多，信息处理时间和价格成本都会增加，从而影响系统的性价比。因此需要认真考量如何优化选择数据源。

2. 支持多字段搜索，可以提高专利检索的准确性，但由于许多扩展字段直接来源于对数

据源的深度加工，因此部分扩展字段的增加可能会影响系统性能。因此在设计专利信息系统时，扩展字段的选择需要考虑系统成本。

3. 社会网络分析功能有较大发展空间，社会网络分析的先进算法^[25]可进一步应用于专利分析领域。

4. 专利评估模块功能由单参数评估功能向多参数评估功能转化已成为必然趋势。如果能开发自动评估算法进行多参数评估，可获得更客观的评估结果。

5. 可视化技术在专利信息系统中被广泛使用，尤其是在分析模块。如 Incopat 中的统计分析功能结合了可视化的分析图和具体的数据表，多层次地呈现了分析结果。Patsnap 中的专利地图结合聚类功能和文本挖掘技术，将复杂的分析结果用图形形式表示，帮助实现更方便直观的分析。可以预计的是，可视化技术将进一步应用于专利信息系统的其他功能模块如评估模块，实现革命性的实时交互和图像化直观分析，

提升用户的工作效率和利用专利资源的能力。

4 面临的问题及挑战

专利信息系统蓬勃发展，并已经进入成熟的使用阶段，但仍存在一些需要解决的问题。

专利仍按元数据进行分类，与内容有关的技术分类和功效分类未获得足够重视。关键字为基础的专利检索，无法实现语义智能识别，检索结果会掺杂与待检目标技术无关的内容。这会导致查全率、查准率和专利资源使用效率的下降，也大大地增加了技术功效矩阵分析时关键词标识的工作量。对此，可能的解决方案是充分引入成熟的文本挖掘技术，实现文本挖掘和自动学习。这将是很有价值的工作。

现有专利信息系统的决策类功能相对偏弱，仅 Thomson Innovation 具备部分专利监测和专利预警功能。因此，决策模块将成为专利信息系统的一个发展重点。特别是建立关键行业的专利信息决策系统（如药业^[26]和精密制造业*），综合专利信息、学术文献和政府报告等多源多维信息，实现强大检索功能，并通过全面技术分析给予精准有效的决策建议。

现有专利信息系统功能多样，但人机交互设计方面考虑不够充分。通常，具备专利工作相当经验的用户才能熟练正确地使用系统的各项功能，初级从业人员和普通用户则会遇到一定困难。因此，未来的专利信息系统需要增加人机交互功能，提升系统友好性，避免因用户使用经验不足所导致的问题。同时由于无线网络技术的发展，专利信息系统也要加强对无线

设备用户的支持。

5 总结

作为科技创新领域成果的重要体现形式之一，专利信息在创新领域中的地位非常重要。由于专利信息系统可以有效提高专利工作者和研发人员的工作效率和创新能力，研究和完善专利信息系统具有重要意义和实用价值。本文阐述了专利信息系统的三个功能模块：专利检索、专利分析和专利评估，对比分析了常见专利信息系统的功能和设计特点，进而提出了目前专利信息系统面临的问题和挑战。通过本文的研究，可以深入了解专利信息系统运行机制、系统架构和具体功能，有利于更好地设计和运用专利信息系统。

参考文献

- [1] 国务院. 国务院关于印发“十三五”国家科技创新规划的通知[EB/OL]. [2017-08-28]. http://www.most.gov.cn/mostinfo/xinxifenlei/gjkjgh/201608/t20160810_127174.htm.
- [2] 王黎莹, 王佳敏, 虞微佳. 区域专利密集型产业创新效率评价及提升路径研究——以浙江省为例[J]. 科研管理, 2017, 38(3): 29-37.
- [3] 科睿唯安信息服务有限公司. Thomson Innovation [DB/OL]. <http://ip-science.thomsonreuters.com.cn>.
- [4] 北京合享智慧科技有限公司. Incopat科技创新情报平台[DB/OL]. <http://www.incoindex.com>.
- [5] 智慧芽信息科技有限公司. PatSnap智慧芽平台[DB/OL]. <https://analytics.zhihuoya.com>.

- [6] 同方知网数字出版技术股份有限公司. 中国知网专利系统[DB/OL]. <http://www.cnki.net>.
- [7] 保定市大为计算机软件开发有限公司. Innojoy专利搜索引擎[DB/OL]. <http://www.innojoy.com>.
- [8] 李杰. 安全科学技术信息检索基础[M]. 北京: 首都经济贸易大学出版社, 2014.
- [9] 李颖, 赵蕴华, 霍翠婷, 等. 日本专利检索体系中主题分类“FI/F-term”的理论与应用研究[J]. 数字图书馆论坛, 2008(11): 11-17.
- [10] 唐炜. 面向战略决策服务的专利分析指标研究[D]. 北京: 中国科学院研究生院(文献情报中心), 2006.
- [11] 汪凯, 张学敏. 基于Incopat的安徽医科大学专利分析[J]. 中国科技信息, 2017(1): 19-21.
- [12] 文庭孝. 专利信息计量研究综述[J]. 情报、信息与共享, 2014(5): 72-80.
- [13] 陈云伟, 方曙. 社会网络分析方法在专利分析中的应用研究进展[J]. 竞争情报, 2012, 56(4): 90-95.
- [14] 许琦, 林晗. 基于专利引证网络的专利产业化评价方法研究—以燃料电池领域为例[J]. 数字图书馆论坛, 2016(9): 49-56.
- [15] 徐丹丹. 专利文本聚类分析及可视化研究[D]. 南京: 南京理工大学, 2009.
- [16] 陈亮, 张静, 张海超, 等. 层次主题模型在技术演化分析上的应用研究[J]. 图书情报工作, 2017(5): 103-108.
- [17] Huang PF, Wei CG, Pan RN. Patent Map Analysis of Measurement Technology[J]. Key Engineering Materials, 2016(680): 60-67.
- [18] 韩红旗, 安小米, 朱东华, 等. 专利技术术语共现的战略图分析方法[J]. 计算机应用研究, 2011, 28(2): 576-579.
- [19] Javadi S, Hashemy SM, Mohammadi K. Classification of Aquifer Vulnerability Using K-means Cluster Analysis[J]. Journal of Hydrology, 2017(549): 27-37.
- [20] 王祥瑞. 大规模数据集高效数据挖掘算法研究[J]. 信息技术与信息化, 2015(9): 256-258.
- [21] 张兆锋. 专利技术功效图应用模式研究[J]. 数字图书馆论坛, 2016(6): 34-39.
- [22] 胡菊香, 吕学强, 刘秀香, 等. 专利技术功效短语获取研究[J]. 科学技术与工程, 2016, 16(14): 227-235.
- [23] 黄继东. 同族专利及其检索[J]. 情报科学, 2002, 20(7): 717-718.
- [24] 科瑞唯安公司. 德温特数据库[DB/OL]. <https://clarivate.com/products/derwent-world-patents-index>.
- [25] 马敏. 社会网络中节点排名与结构相似度量算法研究[D]. 兰州: 兰州大学, 2015.
- [26] 北京东方灵盾科技有限公司. 世界传统药物专利数据库[DB/OL]. <http://www.eastlinden.com>.