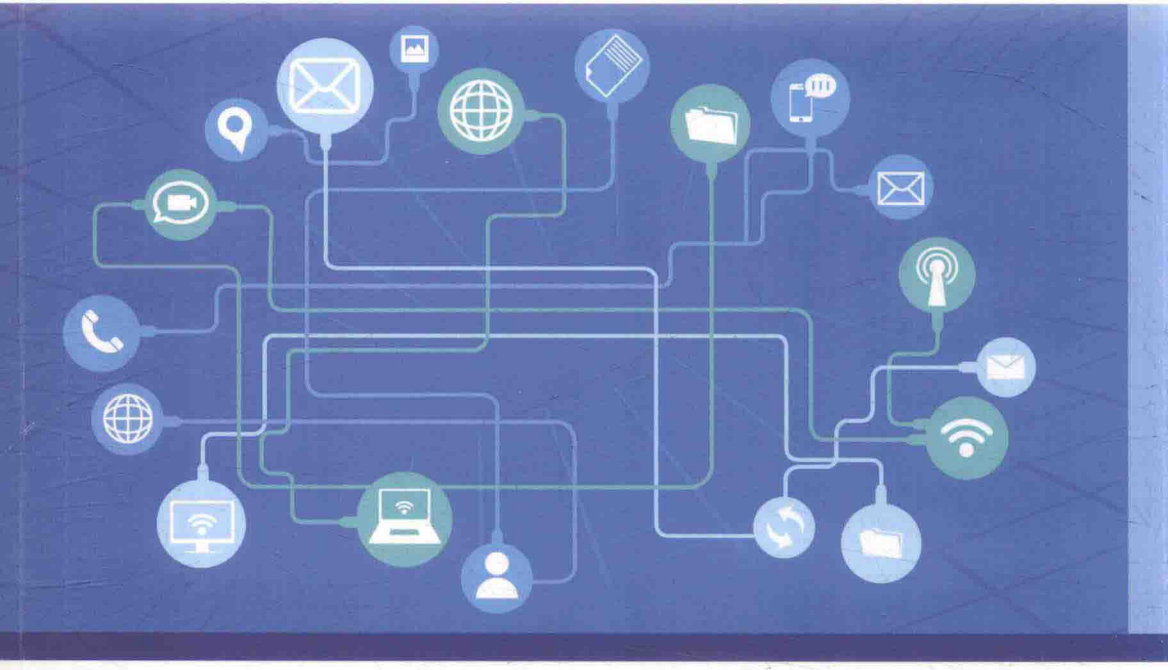
专利信息资源挖掘与发现关键技术研究

刘　耀　朱礼军　靳　玮◎著



科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS

国家科技支撑计划课题“专利信息资源挖掘与发现关键技术研究”
（2013BAH21B02）特别资助

专利信息资源挖掘与发现关键技术研究

刘　耀　朱礼军　靳　玮　著

科学技术文献出版社
SCIENTIFIC AND TECHNICAL DOCUMENTATION PRESS
·北京·

图书在版编目（CIP）数据

专利信息资源挖掘与发现关键技术研究 / 刘耀，朱礼军，靳玮著. —北京：科学技术文献出版社，2018.2
ISBN 978-7-5189-3616-8

Ⅰ.①专… Ⅱ.①刘… ②朱… ③靳… Ⅲ.①专利—信息资源管理—研究 Ⅳ.① G306.3

中国版本图书馆 CIP 数据核字（2017）第 284114 号

专利信息资源挖掘与发现关键技术研究

策划编辑：周国臻　责任编辑：周国臻　白建刚　责任校对：文　浩　责任出版：张志平

出 版 者　科学技术文献出版社
地　　址　北京市复兴路15号　邮编 100038
编 务 部　(010) 58882938，58882087（传真）
发 行 部　(010) 58882868，58882870（传真）
邮 购 部　(010) 58882873
官方网址　www.stdp.com.cn
发 行 者　科学技术文献出版社发行　全国各地新华书店经销
印 刷 者　北京教图印刷有限公司
版　　次　2018 年 2 月第 1 版　2018 年 2 月第 1 次印刷
开　　本　710×1000　1/16
字　　数　189千
印　　张　12.25　彩插4面
书　　号　ISBN 978-7-5189-3616-8
定　　价　58.00元

前　言

专利文献是一种承载海量关键技术的信息资源。根据世界知识产权组织（WIPO）发布的报告显示，自2012年起中国已成为全世界发明专利申请数量第一大国。在如此庞大的数据量下，如何将专利文献的关键信息呈献给用户，使用户能快速掌握信息要点，是优化专利服务的关键科学问题。

本书是国家“十二五”科技支撑计划课题“专利信息资源挖掘与发现关键技术研究”（课题编号：2013BAH21B02）研究成果的组成部分。课题研究的目标是研究数字化、网络化环境下专利信息资源的知识化组织、高效率存储、智能化分析和可视化展示等关键技术并开发系列软件工具，形成面向技术创新的专利信息资源挖掘与发现的关键技术支撑。

围绕以上目标，课题组通过以国家和企业斥巨资建成的海量专利信息资源和服务体系为基础，在建设专利深度标注与知识抽取的资源基础上，研究了专利文本内容自动标注与集成、专利关联知识网络构建和领域专利知识组织工具等专利信息深度分析与处理技术，开发相关软件工具集，并搭建应用服务平台，以检索、分析、可视化的方式面向行业领域提供专利信息资源挖掘与发现服务。

在课题开展的过程中，课题组注入了夜以继日、砥砺深耕的心血。这也是一个不断思考、实践、提升的过程。我们将研究过程中所知所学所获整理为一本书，希望能为相关读者提供一些可供参考、讨论的方法与实例。

本书内容主要介绍专利信息挖掘与发现涉及的框架、方法与技术研究等内容，包括：

①面向智能信息服务的专利信息组织与自动处理研究。研究专利文本内容自动标注与集成技术、专利关联知识网络构建和领域专利知识组织工具构建等关键技术。

②专利信息分析评价模型及应用模型研究。采用定性分析、定量分析、拟定量分析结合的方式，构建宏观、中观、微观3个层面的专利分析模型，并从多个表现特征入手构建了包含引证分析、技术功效矩阵分析、技术成熟度分析等的预测方法和体系。

③专利信息融合分析研究。整合专利文献中的申请人、法律状态、同族专利、引证文献等关键信息，研究专利信息动态关联分析关键技术。

④专利信息智能检索与语义导航研究。研究基于概念检索的海量专利数据智能搜索引擎、多维索引体系的构建、交互性立体式检索结果揭示、个性化推荐等关键技术。

⑤专利信息可视化研究。研究支持专利语义内容关联分析的专利信息可视化关键技术。

本书在编写过程中，参阅了大量的图书和文献，汲取了很多精髓，特别是引用了部分图表、数据等，在此向有关作者表示诚挚的感谢。在研究过程中也得到诸多业界专家、领导、同事的帮助和支持，在此一并致谢。同时，也向课题组成员表示感谢。

由于书中涉及内容广泛，加之笔者水平有限，书中若有不足、错谬之处，恳请同行专家和读者批评指正，以便再版时进一步修订完善。

刘耀

2018年1月

Contents 目 录

图表目录

1 绪 论

1.1 研究背景与意义

伴随科学技术的迅猛发展，新兴学科、边缘学科不断涌现，使得包括专利文献在内的各种科技信息资源数量急剧增加。缺失专利文献的分析意味着无法回答科技发展过程中的某些重大问题，意味着难以使面向科技创新与科技管理的相关政策达到理想的效果，也难以对其进行测度与评价。目前，科研项目管理及创新主体科研实践过程中，专利信息服务与决策支持能力不够深入和系统。因此，通过开展专利信息资源挖掘与发现关键技术研究，在整合多种资源、工具、技术的基础上，进行专利语义知识内容的自动抽取技术及工具的研发，突破海量知识数据的存储、组织、索引和检索的效率与性能，构建专利知识关联网络，并以检索、可视化的方式进行信息展示，旨在为专利信息挖掘与发现相关技术的研发及关键技术在热点监测评价、专利预警、科研项目管理和重点领域的应用提供有益参考，以促进科技信息、技术资源的传播与交流，提升专利信息服务的社会及行业影响力，吸引企业、科研机构等创新主体对专利信息的关注与实践应用，实现创新资源的优化配置、合理共享与高效利用，提升创新能力与质量。

1.2 研究思路

本书的主要研究目标是研究数字化、网络化环境下专利信息资源的知识化组织、高效率存储、智能化分析和可视化展示等关键技术并开发系列软件工具，形成面向技术创新的专利信息资源挖掘与发现的关键技术支撑。

在建设专利深度标注与知识抽取的知识资源基础上，研究专利文本内容自动标注与集成技术、专利关联知识网络构建技术和领域专利知识组织工具等专利信息深度分析与处理的技术并开发相关软件工具集。结合专利评估与监测的理论研究与模型构建，以及上述专利自动标注与知识抽取的基础资源与工具集的建设等研究成果，进行专利信息融合分析核心技术、专利信息智能检索与语义导航关键技术研发、专利信息可视化技术研发的开发工作。