



龙脉图书

圖書館學 情報學 理論與實踐系列叢書

辛卯孟夏 題

文献信息检索、 分析与应用

WENXIAN XINXI JIANSUO,
FENXI YU YINGYONG

陈振标 著



海坪出版社

文献信息检索、分析与应用

陈振标 著

海洋出版社

2016年·北京

图书在版编目 (CIP) 数据

文献信息检索分析与应用/陈振标著. —北京: 海洋出版社, 2016. 6
ISBN 978 - 7 - 5027 - 9366 - 1

I. ①文… II. ①陈… III. ①情报检索 IV. ①G252.7

中国版本图书馆 CIP 数据核字 (2016) 第 094202 号

责任编辑: 杨海萍 张 欣

责任印制: 赵麟芬

海洋出版社 出版发行

<http://www.oceanpress.com.cn>

北京市海淀区大慧寺路 8 号 邮编: 100081

北京朝阳印刷厂有限责任公司印刷 新华书店发行所经销

2016 年 6 月第 1 版 2016 年 6 月北京第 1 次印刷

开本: 787mm × 1092mm 1/16 印张: 21.75

字数: 379 千字 定价: 32.00 元

发行部: 62132549 邮购部: 68038093 总编室: 62114335

海洋版图书印、装错误可随时退换

前 言

据不完全统计,国内关于文献信息检索书籍不下 600 种,面向对象主要是高校学生。从内容上看,这些书籍大同小异,以传授文献获取途径与方法为核心,大部分篇幅用于介绍各类检索工具及其基本使用方法,所用检索示例大多较为简单,缺乏从实践应用角度归纳出的检索案例,导致许多学生仅把检索工具,特别是数据库,当成在校学习时获取文献的途径,而非一种普遍性获取文献信息,应用文献信息的途径。另外,大部分文献信息检索书籍仅限于编者所在高校内使用,因此,往往把数据库环境设置为校内网,导致学生毕业后不知从何处检索文献信息。

鉴于此,本书旨在商榷为理工科学生、科研人员利用现有检索工具检索、获取文献,以及通过检索工具获取有用信息(情报)的方法,从整体上侧重检索工具的情报学应用。在内容组织上,本书第一章为文献信息检索的基础知识;第二至六章为搜索引擎、学术文献、引文文献、专利文献等常用文献检索工具,以及标准、产品样本、集成电路布图设计、商标、指数等特种文献信息检索工具的介绍,每个章节均选用可反映检索系统功能和特色的案例进行示范,这些案例均来自于笔者所从事的文献情报服务活动,侧重于检索工具的深度检索和实践性应用。第七章着重介绍文献信息检索在科学研究、新产品研制与上市、人才引进、产业现状调研等实践中的应用,并以实际案例详细阐述文献检索在发掘经典、高水平和前沿文献,以及技术方案筛选、产品先进性分析、专利侵权规避、人才评估、产业调查分析中的应用。第八章和第九章侧重文献信息可视化分析方法及实践应用,用实际案例详细阐述可视化分析在挖掘核心研究机构(专利权人)、研究者(发明人),研究产业技术生命周期、关键技术发展趋势,以及产业竞争环境与竞争态势等领域的应用。

本书由五位作者共同完成,每位作者分工如下:

陈振标:负责统筹全书,提出全书编撰思路,并完成第一章第三、四节,第五章第一、二、三节,第六章第五节,第七章第一、二、四节,第八章第一、三、四节及第五节案例 3 内容的编撰任务。

王征清：完成第四章，第六章第三、四节，第七章第三节内容的编撰任务。

陈建芳：完成第三章，第五章第四节内容的编撰任务。

王玲艳：完成第一章第一、二节，第二章，第六章第一、二节内容的编撰任务。

陈慧琪：完成第八章第二节及第五节案例1、案例2内容的编撰任务。

本书由陈振标统稿，陈建芳初校稿，陈振标总校稿。在本书的编写过程中得到了相关老师、朋友的热情帮助，在此深表谢意。特别感谢发起本书编写的黄元石老师，以及为推动本书顺利完成和出版，不辞辛苦、来回奔波，不断创造优良条件的刘敏榕老师。衷心感谢我的爱人，正是她的不断督促和全心照顾孩子，为我编写本书提供了良好时间保证；感谢给我无私的爱和不断鼓励与支持的家人。

由于作者水平有限、时间仓促，书中难免有疏漏和不足，敬请读者谅解，并欢迎批评指正，以便我们共同努力，继续研究和推动文献情报服务向更广更深的方向发展。

编者

2015年9月30日

目 录

第一章 文献与检索	(1)
第一节 信息与文献	(1)
第二节 文献类型与识别	(1)
一、文献类型	(1)
二、常用文献及识别	(3)
第三节 文献信息检索	(19)
一、文献信息检索概念	(19)
二、计算机检索技术	(20)
三、文献信息检索策略	(26)
四、文献信息检索效果评价	(29)
第四节 文献信息检索应用	(30)
一、文献调研检索	(30)
二、科技查新检索	(30)
三、论文收录引用检索	(30)
四、专利事务检索	(31)
五、技术贸易检索	(32)
六、产业调研/分析检索	(32)
第二章 搜索引擎与应用	(33)
第一节 搜索引擎概述	(33)
一、搜索引擎原理	(33)
二、搜索引擎分类	(33)
第二节 谷歌检索方法及特色应用	(34)
一、谷歌高级检索方法	(34)
二、谷歌学术搜索	(39)
三、谷歌图书搜索	(44)
四、谷歌专利搜索	(46)
第三节 百度特色检索应用	(48)

一、百度指数	(49)
二、百度识图	(53)
第三章 科技文献检索与应用	(56)
第一节 中国知网(CNKI)	(56)
一、检索方法	(56)
二、检索应用案例	(69)
第二节 万方数据知识服务平台	(73)
一、检索方法	(74)
二、检索应用案例	(81)
第三节 维普中文科技期刊数据库	(83)
一、检索方法	(84)
二、检索应用案例	(86)
第四节 读秀学术搜索系统	(88)
一、检索方法	(88)
二、检索应用案例	(96)
三、CNKI、万方、维普、读秀应用对比	(101)
第四章 引文文献检索与应用	(103)
第一节 引文文献检索系统概述	(103)
一、科学引文索引 SCI	(103)
二、社会科学引文索引 SSCI	(103)
三、艺术与人文引文索引 A&HCI	(104)
四、科技会议录索引 CPCI-S 引文检索系统	(104)
五、Scopus 引文检索系统	(105)
六、中国科学引文索引数据库 CSCD	(105)
七、中文社会科学引文索引 CSSCI	(106)
八、中国引文数据库 CCD	(106)
第二节 中国引文数据库 CCD	(107)
一、检索方法	(107)
二、检索应用案例	(114)
第三节 科学引文索引扩展版(SCIE)	(117)
一、检索规则	(117)
二、检索方法	(118)
三、检索应用案例	(128)

第四节 SCOPUS 数据库	(131)
一、检索规则	(131)
二、检索方法	(132)
三、检索应用案例	(135)
第五章 专利文献检索与应用	(138)
第一节 专利及专利文献	(138)
一、专利及专利文献	(138)
二、专利授予原则	(138)
三、专利权特点	(139)
四、专利分类法	(140)
五、专利应用价值与风险	(141)
第二节 专利文献检索系统	(142)
一、专利行政机构提供的免费专利数据库	(142)
二、商业性专利数据库	(145)
三、其它专利数据库	(149)
第三节 专利之星(PATENTSTAR)专利检索系统	(150)
一、专利之星概述	(150)
二、中国专利检索	(152)
三、世界专利检索	(168)
四、检索应用案例	(170)
第四节 外观设计专利设计检索系统	(174)
一、外观设计专利及其检索	(174)
二、家具行业外观设计专利图像检索服务平台	(174)
三、检索应用实例	(179)
第六章 特种文献检索与应用	(183)
第一节 产品样本检索	(183)
一、产品样本概述	(183)
二、尚唯产品样本数据库(GPD)	(183)
第二节 标准文献检索	(188)
一、国家标准馆概述	(188)
二、国家标准文献共享服务平台(CSSN)	(188)
第三节 集成电路布图设计检索	(194)
一、集成电路布图设计概述	(194)

二、集成电路布图设计专有权检索平台	(195)
第四节 商标检索	(199)
一、商标概述	(199)
二、商标网上查询系统	(199)
第五节 指数工具应用	(207)
一、指数概述	(207)
二、CNKI 指数	(207)
三、万方知识脉络分析	(212)
四、指数应用案例	(213)
第七章 文献信息应用	(219)
第一节 科学研究中的文献信息应用	(219)
一、科学研究与文献信息应用	(219)
二、科学研究中文献信息应用的策略与步骤	(220)
三、案例:创业板研究中的文献调研	(221)
第二节 新产品研制与上市中的专利信息应用	(232)
一、产品研制与上市中需要解决的主要问题	(232)
二、产品研制与上市中的专利信息运用策略与步骤	(232)
三、案例 1:技术方案筛选中的专利信息应用	(234)
四、案例 2:产品先进性分析中的专利信息应用	(237)
五、案例 3:主动侵权检索中的专利信息应用	(239)
六、案例 4:被控侵权检索中的专利信息应用	(242)
第三节 人才引进中的文献信息应用	(242)
一、人才引进中的查收查引需求	(242)
二、人才引进中的文献信息应用基本步骤	(243)
三、案例 1:利用 CCD 评价人才科研能力	(243)
四、案例 2:利用 SCIE 数据库评价人才科研能力	(247)
第四节 产业调研中的文献信息应用	(252)
一、产业调研中的文献信息需求	(252)
二、产业调研中的文献信息运用策略与步骤	(252)
三、案例:福建省集成电路产业现状调研	(253)
第八章 文献信息可视化分析	(267)
第一节 可视化概述	(267)
一、科学计算可视化	(267)

二、信息可视化	(267)
第二节 科技文献信息可视化分析	(270)
一、词频分析(Word Frequency Analysis)	(270)
二、引文分析(Citation Analysis)	(270)
三、共引分析(Co-citation Analysis)	(270)
四、共词分析(Co-word Analysis)	(271)
五、聚类分析(Cluster Analysis)	(271)
六、社会网络分析(SocialNetworkAnalysis, SNA)	(271)
第三节 专利文献信息可视化分析	(272)
一、专利定量分析方法	(272)
二、专利定性分析方法	(277)
三、专利拟定量分析方法	(277)
第四节 文献信息可视化分析工具	(278)
一、科技文献信息可视化分析工具	(278)
二、专利文献信息可视化分析工具	(282)
第五节 文献信息可视化分析应用案例	(286)
案例1 基于 CITESPACE II 的 IGZO 科技文献分析	(287)
案例2 基于 CITESPACE II 的 IGZO 专利文献分析	(305)
案例3 基于 PIAS 的 OLED 专利文献分析	(316)
参考文献	(333)

第一章 文献与检索

第一节 信息与文献

1. 信息

信息是指一切事物自身存在方式以及它们之间相互关系、相互作用等运动状态的表达。信息资源是指各种载体和形式的信息的集合,包括文字、音像、印刷品、电子信息、数据库等。

2. 文献

“文献”一词最早见于《论语·八佾》,南宋朱熹《四书章句集注》认为“文,典籍也;献,贤也”,即,文指的是典籍文章,献指的是古代先贤的见闻、言论以及他们所熟悉的各种礼仪和自己的经历。1983年公布的中华人民共和国国家标准《文献著录总则》(GB/T 3792.1-1983)将“文献”定义为:记录有知识的一切载体。根据该定义,除书籍、期刊等出版物外,凡载有文字的甲骨、金石、简帛、拓本、图谱乃至缩微胶片、视盘、声像资料等等,皆属文献的范畴。

3. 信息、文献之间的关系

文献和信息是从不同角度对同一种事物的表述:文献侧重于载体的概念,强调其作为载体记录人类客观知识的价值;而信息侧重效用性,无论其载体形式如何,强调其作为一种智力资源的开发与利用。本书在后续章节中统一将其称为文献信息。

第二节 文献类型与识别

一、文献类型

文献信息种类繁多,一般可按载体形式、出版类型、加工层次等方式进

行分类。

（一）按物质载体形式进行划分

1. 印刷型：是文献的最基本方式，包括铅印、油印、胶印、石印等各种资料。其优点是可直接阅读，使用方便；其缺点是存储密度低、收藏占用空间大。

2. 缩微型：是以感光材料为载体的文献，又可分为缩微胶卷和缩微平片。其优点是存储密度大、体积小、便于保存、转移和传递。缺点是阅读时须用阅读器。

3. 机读型：是利用计算机进行存储和阅读的一种文献信息形式。它通过编码和程序设计，把文献变成符号和机器语言，输入计算机，存储在磁带或磁盘上，阅读时，再由计算机输出。其优点是它能存储大量信息，可按多种形式组织这些信息，并能以极快的速度从中取出所需信息。

4. 声像型：又称直感型或视听型，是以声音和图像形式记录在载体上的文献，如唱片、录音带、录像带、科技电影、幻灯片等。

（二）按文献出版类型划分

按出版类型的不同，文献可分为图书、期刊、特种文献三大类。其中，特种文献包括：专利文献、标准文献、学位论文、科技报告、会议文献、政府出版物、档案资料、产品资料等。

（三）按文献内容加工层次划分

1. 零次文献：指未经正式发表或未形成正规载体的一种文献形式。如：书信，手稿，会议记录，笔记等。零次文献在原始文献的保存、原始数据的核对、原始构思的核定（权利人）等方面有着重要的作用。

2. 一次文献：又称原始文献，指以作者本人的工作经验、观察或者实际研究成果为依据而创作的具有一定新见解的原始文献。如，期刊论文、研究报告、专利说明书、会议论文、学位论文、技术标准等。

3. 二次文献：是指文献工作者对一次文献进行加工、提炼和压缩之后所得到的产物，即对无序的一次文献的外部特征如题名、作者、出处等进行著录，将其内容压缩成简介、提要或文摘，并按照一定的学科或专业加以有序化而形成的文献形式，如目录、题录、文摘杂志（包括简介式检索刊物）等。二次文献都可作为文献检索工具，能比较全面、系统地反映某个学科、专业或专题在一定时空范围内的文献线索，是积累、报道和检索文献资料的有效手段。检索工具书和网上搜索引擎都是典型的二次文献。

4. 三次文献：是按给定的课题，利用二次文献选择有关的一次文献加以分析、综合而编写出来的专题报告或专著。如，综述报告、专题述评、学科年度总结、进展报告、研究报告、技术预测、数据手册等。人们常把这类文献称为“情报研究”的成果。

二、常用文献及识别

（一）图书

凡由出版社（商）出版的不包括封面和封底在内 49 页以上的印刷品，具有特定的书名和著者名，编有国际标准书号，有定价并取得版权保护的出版物称为图书。

图书的内容比较系统、全面、成熟、可靠，但出版周期较长，传递信息速度较慢。

1. 图书特征

在形式上，图书在其封面上具有书名、责任者、出版社等信息，在扉页中有出版时间、ISBN、版次、印刷时间、页数、字数等信息，如图 1.2-1。



图 1.2-1 图书

ISBN 号是识别图书的最主要依据。如果是馆藏图书，一般还有索书号。

（1）ISBN

国际标准书号（International Standard Book Number，简称 ISBN）是由 EAN、UCC 前缀、组区号、出版者号、出版序号、校验码四部分共 13 个数字组成。前缀与数字之间有半个汉字宽的间隔，数字间用连字符隔开。其中：

EAN. UCC 前缀: 由三位数字组成, 是国际物品编码协会分配的产品标识编码。图书的 EAN. UCC 前缀为 978。

组区号: 代表一个语言或地理区域, 国家或集团的代码, 由国际标准书号中心分配。取值范围为: 0 - 7, 80 - 94, 950 - 995, 9960 - 9989, 99900 - 99999, 共 199 个, 按出版量愈大, 组号愈短的原则分配。我国大陆地区组区号为 7, 香港为 962, 澳门为 972, 台湾为 957。

出版者号: 由组区或国家 ISBN 中心分配, 并按出版社出版量愈大, 出版者号愈短的原则分配。其中非连续出版的电子出版物的出版者号均为 6 位数。

出版序号: 由出版社自行分配, 一个图书一个号码, 也称为书名号。出版者号和出版序号连在一起共为 8 位数字。

校验号: 固定用一位数字表示。采用模数 10 的加权算法对前 12 位数字计算后得出。

如: ISBN978 - 7 - 115 - 23158 - 1, 978 是 EAN. UCC 前缀, 表示该商品是图书; 7 代表中国大陆出版的图书; 115 表示人民邮电出版社; 23158 则是人民邮电出版社给该书的出版序号; 1 为校验号。

国际标准书号的使用范围包括: 教科书、印刷品、缩微制品、教育电视或电影、混合媒体出版物、微机软件、地图集和地图、盲文出版物、电子出版物。

(2) 索书号

索书号是图书馆藏书排架用的编码, 又称索取号。索书号是文献外借和馆藏清点的主要依据, 也是向图书馆借阅图书的最重要线索。

索书号一般由分行排列的几组号码组成, 一个索书号代表一种书, 常被印在目录卡片的左上角、书脊下方的书标上以及图书书名页或封底的上方。按图书馆藏书排架方法的不同, 索书号可分为分类索书号和形式索书号两大类。

通常情况下, 索书号由两部分组成: 第一部分是分类号, 即根据图书的学科主题所取用的分类号码; 第二部分是书次号, 即按照图书作者姓名所编排的著者号码, 或者是按照图书进入馆藏时间的先后所取用的顺序号码。

如: 王小宁主编的《电子商务物流管理》索书号为: F713.365.1/884。其中, F713.365.1 是中图分类号, 884 是著者号。

注: 著者号根据著者姓名拼音字母编排得到, 其对应顺序如表 1.2 - 1。

表 1.2-1 著者号编排规则

拼音字母	对应数字	拼音字母	对应数字
ABC	0	PQ	5
DEF	1	RS	6
GHI	2	TU	7
JKL	3	WX	8
MNO	4	YZ	9

(3) 图书分类号

图书分类法又叫图书分类词表,是按照图书的内容、形式、体裁和读者用途等,在一定的哲学思想指导下,运用知识分类的原理,采用逻辑方法,将所有学科的图书按其学科内容分成几大类,每一大类下分许多小类,每一小类下再分子小类。最后,每一种书都可以分到某一个类目下,每一个类目都有一个类号。《中国图书馆分类法》(简称《中图法》)和《中国科学院图书馆图书分类法》(简称《科图法》)是国内常用的图书分类法。

《中国图书馆分类法》(原称《中国图书馆图书分类法》)是我国建国后编制出版的一部具有代表性的大型综合性分类法,是当今国内图书馆使用最广泛的分类法体系。《中图法》初版于1975年,2010年出版了第五版。第五版对与人类生活息息相关的经济、生产和生活服务业(包括金融、房地产、公共设施、社会福利、娱乐业等),以及发展迅速的通信业、交通运输业、计算机技术等方面的大类进行了重点修订,使其更符合社会发展趋势。目前,《中图法》已普遍应用于全国各类型的图书馆,国内主要大型书目、检索刊物及各类型机读数据库等都著录《中图法》分类号。

《中图法》使用字母与数字相结合的混合号码,分为五大部类,二十二个大类,51 881个类目(包括通用类目),如表1.2-2。

表 1.2-2 中图法

五大部类	22 个大类
第一部类	A 马克思主义、列宁主义 毛泽东思想、邓小平理论
第二部类	B 哲学、宗教

续表

五大部类	22 个大类
第三部类 (社会科学)	C 社会科学总论 D 政治法律 E 军事 F 经济 G 文化、科学、教育、体育 H 语言、文字 I 文学 J 艺术 K 历史、地理
第四部类 (自然科学)	N 自然科学总论 O 数理科学和化学: 01 数学 03 力学 04 物理学 06 化学 07 晶体学 P 天文学、地球科学 Q 生物科学 R 医药、卫生 S 农业科学 T 工业技术; U 交通运输 V 航空、航天; X 环境科学安全科学
第五部类	Z 综合性图书

2. 图书种类

图书按其用途可分为以下 3 种类型:

①阅读用书: 包括教科书、专著、文集等。

②参考工具书: 包括字典、词典、百科全书、年鉴、手册、名录、图录、年表、历表等, 是供人们释疑解难、翻检查考的出版物。

③检索用书: 以图书形式刊行的书目、题录、文摘等, 供人们查找一定范围内信息线索的出版物。

3. 图书检索入口

图书检索入口一般有: 书名、作者、主题词、ISBN、出版社、出版日期、丛书名、中图法分类号、索书号, 部分检索系统还提供目次、全文检索入口。

(二) 期刊

期刊, 也称杂志, 是指有固定名称、每期版式基本相同、定期或不定期

的连续出版物。它的内容一般是围绕某一主题、某一学科或某一研究对象，由多位作者的多篇文章编辑而成，用卷、期或年、月顺序编号出版。

1. 期刊特征

期刊具有如下特点：①定期连续出版，有出版序号；②有长期固定统一的刊名，开本、篇幅、栏目等基本稳定；③内容新颖，时效性强；④信息量大，作者众多。

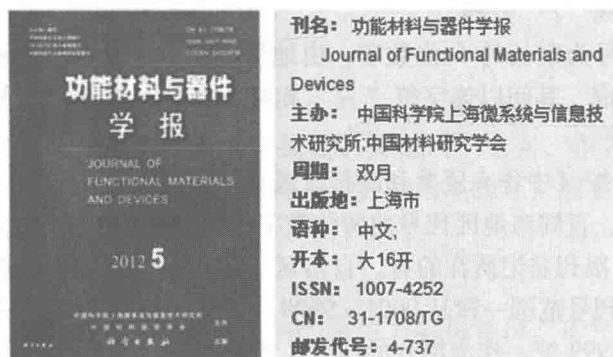


图 1.2-2 期刊

从形式上，识别期刊的最主要特征是期刊具有国际标准期刊号（ISSN）。如果是国内刊物，还有国内统一刊号（CN）。

（1）国际标准期刊号（ISSN）

国际标准期刊号（International Standard Serial Number，简称 ISSN），也称国际标准连续出版物编号，是根据国际标准 ISO3297 制定的连续出版物国际标准编码，其目的是使世界上每一种不同题名、不同版本的连续出版物都有一个国际性的唯一代码标识。

ISSN 由设在法国巴黎的国际 ISDS 中心管理。该编号是以 ISSN 为前缀，由 8 位数字组成。8 位数字分为前后两段各 4 位，中间用连接号相连，格式如下：ISSN XXXX - XXXX，如：Science（Print ISSN 0036 - 8075；online ISSN 1095 - 9203），前 7 位数字为顺序号，最后一位是校验位。ISSN 通常都印在期刊的封面或版权页上。

1985 年我国政府正式批准 ISSN 中国国家中心成立，中国国家中心设在国家图书馆内。ISSN 中国国家中心负责经国家新闻出版署正式批准出版的连续出版物 ISSN 的分配、管理、使用和咨询，中国连续出版物书目数据送交 ISSN