

化学及相关学科信息源

信息检索、分析与管理

冯白云 林佳 编著



清华大学出版社

化学及相关学科信息源

—— 信息检索、分析与管理

冯白云 林佳 编著

清华大学出版社
北京

内 容 简 介

信息源是实施检索的基础,当今各类电子型信息源不仅具备日益完善的检索功能,而且拥有多方位的信息分析功能,并提供了组织管理信息的各种模式和工具,所以,本书定名为《化学及相关学科信息源——信息检索、分析与管理》。本书以信息源为本,紧密联系信息源,全面系统地阐述信息检索、检索结果分析评估及信息成果组织管理的原理和方法。它们是科学活动的重要组织部分,是创新的基础,也是科研人员必须具备的基本素质和能力,即信息素养(information literacy)。

全书共12章。第5至第11章为与化学及相关学科有关的主要信息源的剖析和介绍,原理、思路和方法寄寓于其中;第1至第4章、第12章则概括地阐述了相关领域的基础知识,是原理、思路和方法的深化。信息资源电子化、运行网络化是基本信息环境,所以本书以计算机检索、分析和组织管理为主,传统的手工方式只作为比较和参照。

本书为化学及相关学科(包括化工、环境、生物、材料等)信息素养教育的基础课程教材,也可以作为相关领域专业人员及图书馆人员的参考用书。

图书在版编目(CIP)数据

化学及相关学科信息源——信息检索、分析与管理/冯白云,林佳编著. —北京:清华大学出版社,2003

ISBN 7-302-06383-4

I. 化… II. ①冯… ②林… III. ①化学—情报检索 ②化学—信息—分析
③化学—信息管理 IV. G252.7

中国版本图书馆 CIP 数据核字(2003)第 015841 号

出 版 者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

责任编辑: 刘明华

印 刷 者: 北京鑫海金澳胶印有限公司

发 行 者: 新华书店总店北京发行所

开 本: 787×960 1/16 印张: 20.75 字数: 380 千字

版 次: 2003 年 5 月第 1 版 2003 年 5 月第 1 次印刷

书 号: ISBN 7-302-06383-4/G·318

印 数: 0001~5000

定 价: 26.00 元

前 言

1991年,冯白云、李京华编著的《化学化工文献检索与利用》,先后印行两次,受到学生和其他读者的普遍欢迎和好评,荣获清华大学优秀教材二等奖。然而,日月沧桑,十多年来信息技术发生了翻天覆地的变化;当年,计算机检索尚局限于联机检索;当今,联机检索、光盘检索、网络检索……实现了信息资源电子化、多元化,运行环境网络化的巨大飞跃。显然,原来那本以印刷型信息源、手工检索为主体,只涉及国际联机检索的《化学化工文献检索与利用》已完全过时。因此,本书是其必然发展。

同时,作者分析了近期出版的相关著作,虽然增设了光盘检索、网络检索的章节,但是当属锦上添花,未改变陈旧的体例、格局和内容。传统过时的手工检索工具及其检索方法仍占这类书的85%左右,因此,也已完全不适用。这是编著本书的动机之一。

近年来,随着信息技术的发展,涉及计算机检索主题的著作像雨后春笋,相继应运而生,而且不乏好书。如《Internet 网络信息检索——原理 工具 技巧》(储荷婷、张晓林、王芳,清华大学出版社,1999年)。它们对普及、推动信息检索知识、技术的应用和强化信息素养教育,起到了积极的作用。然而,信息检索原理、方法和技术是与信息源紧密相关的,而且化学及相关学科领域(包括化工、环境、生物、材料等)的信息源极其丰富,具有极深的学科特殊性。必须通晓上述领域信息源,掌握其特殊的检索语言、检索入口等方面的学科特点,才能真正掌握检索原理、方法和技术,达到得心应手、随心所欲地获取所需学科信息的目的。在这方面,上述著作是无法满足化学及相关领域专业人员需求的。从此意义而论,本书恰好弥补了专业学科需求的空白,因此,编著本书是客观需求的选择。

强化信息素养(information literacy)教育是知识经济、信息时代培养学生自主学习能力和创新能力的重要举措,是素质教育不可或缺的重要方面。美国自20世纪70年代末提出了“信息素养”的概念,2000年年初美国大学图书馆协会制定并批准和公布了“高等教育信息素养能力标准”(Information Literacy Competency Standards for Higher Education)作为推进美国大学信息素养教育的重大举措。近两年来已得到了广泛的响应,成为美国高等教育教学改革的一个热点。我国自2000年3月始在教育学、图书馆学和情报学领域的期刊上已发表了近50篇论述信息素养教育的学术论文。对信息素养教育已形成

了较为一致的认识:信息素养教育是信息时代人们必备的基本素养,它包括信息文化素养、信息意识素养和信息技能素养3个方面。信息素养教育的目标和内容是使人们具备:①正确确定所需信息的性质和内容;②高效检索信息;③正确地分析评价信息和信息源;④组织、管理信息,形成信息知识库;⑤正确地利用信息进行创新,完成相关任务;⑥通晓并自觉遵循信息运行过程中所涉及的伦理道德和相关法律。因此,将培养信息素质作为本书的基本指导思想,使之成为信息素养教育的基础课程教材。基于此,作如下考虑:①以现代信息技术为主,即以电子信息源与网络运行环境下的信息检索原理、方法和技术为主;同时,正确处理与传统的文献知识、印刷型信息源的关系(主要作为对比,有利于深化对电子信息源及计算机检索的理解);②强化检索、评价分析及组织管理信息等3方面能力的训练是本书取材和结构全书的依据;③强调学科特色,强调紧密联系学科最有代表性的、最重要的信息源阐述检索原理、方法和技术是本书的基本思路。

由于信息源是实施检索的对象,通晓信息源,善于分析、比较和取舍信息源,并能恰到好处地对相关信息源实施相互补充,是检索者实施检索的立足点。所以本书以极大篇幅介绍信息源,并强调:①主次分明,一是化学及相关学科最有代表性的,也是最为重要的;二是突出功能、特点及相应原理、方法和技术运用;②以电子信息源为主,若已有网络版的则以网络版为主;③加强参照,使本书第3章和第4章所阐述的基本原理和方法与实际运用紧密联系起来。

本书着重于化学及相关学科的科学文献和数据事实检索,因此因特网上通用的网络资源及网络检索工具非本书范围,相关知识请参考有关著作。

联系实际,深入浅出,强化思路和基本信息素养的训练,即“授人以渔”,是编著者遵循追求的宗旨。

本书是大学化学及相关学科目前所设的信息检索课及即将设置的信息素养教育课程的教学用书;也可作为化学及相关学科专业人员及图书馆、信息部门参考咨询人员的参考书。

本书的编著者之一是清华大学图书馆冯白云教授,出自化学化工专业,从事文献检索课教学和图书馆参考咨询工作20余年,曾任清华大学图书馆学术委员会主任、文献检索教研组主任;编著者之二是林佳副教授、化学硕士,现为化学化工类文献检索课主讲教师。本书共分12章,由冯白云拟订大纲,其中第1~3章、第5章、第8~11章由冯白云撰写,第4章、第6章、第7章、第12章及附录由林佳编著,全书由冯白云统稿并审定,电子书稿的加工处理工作全部由林佳完成。

最后,我们要深切感谢清华大学图书馆文献检索教研组主任孙平教授的关

心、指导,感谢文献检索教研组全体同仁的帮助,感谢清华大学图书馆领导给予的关怀和支持。我们也感谢清华大学出版社的大力支持和帮助,使本书得以顺利出版。

信息技术瞬息万变,由于我们水平所限,本书定有不妥、误漏之处,敬请读者和同行指正。

编著者
2002年4月

目 录

1 绪论	1
1.1 研究对象	1
1.2 基本体系	2
1.3 几个常用术语之基本含义	3
1.3.1 信息(information)	3
1.3.2 知识(knowledge)	5
1.3.3 文献(document, literature)	5
1.3.4 出版物(publication)	5
1.3.5 信息源(information source)	6
1.3.6 信息资源(information resource)	6
1.4 化学及相关学科信息源及其检索的特殊性	6
1.5 自学,在实践中自学	8
2 文献信息源	9
2.1 概述	9
2.2 原始文献(一次文献)(primary document)	9
2.2.1 阅读类图书(reading book)	9
2.2.2 期刊论文(journal article)	10
2.2.3 特种文献(special document)	11
2.3 文摘-索引和书目(或文献数据库和书目数据库)—— 二次文献(second document)	15
2.3.1 文摘	15
2.3.2 索引	16
2.3.3 书目(bibliography, catalogue)	17
2.3.4 结论	18
2.4 参考工具书——三次文献(tertiary document)	18
2.5 文献的介质形式	19
2.5.1 印刷型文献(printed document)	19
2.5.2 缩微型文献(microform)	19

2.5.3	声像型文献(audio-video)	19
2.5.4	机读型文献	20
2.6	电子文献(electronic document)或电子出版物 (electronic publication)及文献集(collection)	20
2.6.1	单元记录(record)	21
2.6.2	索引库(或称词库)	22
2.6.3	检索软件(information retrieval package)	22
2.7	原始文献出版类型识别	23
3	检索原理和方法	31
3.1	文献信息流	31
3.2	检索提问与文献信息源的选择	32
3.2.1	检索提问基本属性分类	33
3.2.2	文献检索所涉及的检索提问分析	33
3.2.3	信息源与提问属性的对应关系	34
3.2.4	小结	35
3.3	检索入口(检索点)与检索途径	35
3.4	检索语言类别及其与检索入口、检索途径的关系	38
3.5	主题语言及词表	39
3.5.1	主题语言类别	39
3.5.2	主题语言发展趋势	40
3.5.3	词表	41
3.6	分类语言及分类法(分类表)	42
3.6.1	分类语言的含义、功能和类型	42
3.6.2	分类表	44
3.7	检索策略	45
3.7.1	概述	45
3.7.2	主题分析	46
3.7.3	检索式及其实施	46
3.7.4	小结	49
3.8	著者和机构名称的表述方法	50
4	计算机信息检索	52
4.1	概述	52

4.1.1	定义	52
4.1.2	计算机信息检索的发展历史	52
4.1.3	计算机信息检索的类型	53
4.2	计算机信息检索系统与数据库	54
4.2.1	概述	54
4.2.2	检索系统的类型	54
4.2.3	数据库的类型	55
4.2.4	数据库的结构	56
4.3	计算机信息检索的方法与技术	57
4.3.1	逻辑运算符(logical operator)	57
4.3.2	位置算符(proximity operator)	58
4.3.3	通配符(wildcard character)——截词和屏蔽	59
4.3.4	禁用词(stop word)	59
4.3.5	检索字段标识符	60
4.3.6	范围检索	61
4.3.7	加权检索	61
4.3.8	检索结果输出	61
4.3.9	菜单式检索	61
4.3.10	命令式检索	62
4.3.11	超文本检索	62
4.3.12	超媒体检索	64
4.3.13	小结	64
4.4	联机信息检索(online-based information retrieval)	65
4.4.1	概述	65
4.4.2	拨号联机检索	65
4.4.3	网络联机检索	65
4.4.4	国际著名联机检索系统	66
4.4.5	国内联机检索系统	67
4.5	光盘信息检索	67
4.5.1	光盘数据库的发展历史	67
4.5.2	光盘的记录、读取原理及光盘的特点	68
4.5.3	光盘检索系统	68
4.5.4	光盘检索软件	68
4.6	光盘数据库检索与网络数据库检索的比较	69

4.7	OPAC 及馆藏资源共享	70
4.7.1	OPAC	70
4.7.2	馆藏资源共享	71
5	美国化学文摘社信息源及其检索	72
5.1	美国化学文摘社及其信息源概况	72
5.1.1	概述	72
5.1.2	美国《化学文摘》的变迁	72
5.1.3	美国化学文摘社信息源的优势	73
5.2	CAS Web Site	75
5.3	SciFinder 和 SciFinder Scholar	76
5.3.1	概述	76
5.3.2	CA (Chemical Abstracts)	77
5.3.3	CA REGISTRY	77
5.3.4	CA REACT	78
5.3.5	MEDLINE 和 OLDMEDLINE	78
5.3.6	SciFinder Scholar 2001 的检索方法和检索功能剖析	78
5.4	STN	89
5.4.1	概述	89
5.4.2	STN on the Web	91
5.4.3	STN Easy	92
5.5	CD-ROM	93
5.5.1	CA on CD	93
5.5.2	第 12,13,14 累积期 CD-ROM(12 CI on CD with Abstracts, 13 CI on CD with Abstracts, 14 CI on CD with Abstracts)	96
5.5.3	CA surveyor (TM)	96
5.5.4	CASSI CD-ROM	97
5.6	印刷型《化学文摘》(CA)	97
5.6.1	文摘刊	97
5.6.2	累积索引中的直接性索引	98
5.6.3	间接性、指导性索引	102
5.6.4	资料来源索引(chemical abstracts service source index, CASSI)及原始文献的索取	106

6 美国 ISI 公司信息源	108
6.1 关于 ISI	108
6.2 ISI Web of Knowledge	109
6.3 Science Citation Index (SCI,科学引文索引)	112
6.3.1 印刷版 SCI	112
6.3.2 光盘版数据库 SCI CDE	113
6.3.3 网络版数据库 Web of Science	118
6.4 ISI Chemistry	132
6.4.1 概述	132
6.4.2 ISI Chemistry 检索方法	132
6.4.3 ISI Chemistry 的结构绘制软件	137
6.4.4 检索实例	140
6.5 Index to Scientific & Technical Proceedings (ISTP,科技会议录索引)	144
6.6 Derwent Innovations Index	145
6.6.1 概述	145
6.6.2 检索方法	145
6.6.3 检索结果利用	148
6.7 BIOSIS Previews	150
6.7.1 BIOSIS Previews 概述	150
6.7.2 BIOSIS Previews 检索方法	150
6.7.3 检索结果	152
6.8 Current Contents Connect	153
6.8.1 概述	153
6.8.2 检索方法	154
6.8.3 检索结果	158
6.8.4 Email 通告(Alert)服务	160
6.9 ISI Journal Citation Reports(JCR)	161
6.9.1 概述	161
6.9.2 JCR 常用术语及重要指数	162
7 美国工程信息公司信息源	164
7.1 美国工程信息公司简介	164
7.2 Ei Engineering Index(印刷版《工程索引》)	165

7.2.1	概述	165
7.2.2	Ei Thesaurus(Ei 主题词表)	165
7.2.3	Engineering Index Monthly(工程索引月刊)	167
7.2.4	Engineering Index Annual(工程索引年刊)	167
7.2.5	Engineering Index Cumulative Index (工程索引多年累积索引)	167
7.2.6	PIE (Publications in Engineering, 工程索引来源出版物)	167
7.3	Ei Compendex 数据库	168
7.3.1	概述	168
7.3.2	光盘版数据库 Ei Compendex	168
7.3.3	网络版数据库 Ei CompendexWeb	178
7.3.4	Compendex 数据库的光盘版与网络版比较	187
7.4	信息村	188
7.4.1	工程信息村	188
7.4.2	化学信息村(ChemVillage)	196
8	特种文献信息源	198
8.1	科技报告信息源	198
8.1.1	NTIS 网站检索(NTIS Web Site Search)	200
8.1.2	镜像点网络检索 NTIS 的文献	201
8.1.3	NTIS 光盘数据库	206
8.1.4	《政府报告通报和索引》	206
8.1.5	联机检索数据库	207
8.2	专利文献信息源	208
8.2.1	专利文献基本知识	208
8.2.2	专利文献信息源及其检索	215
8.3	学术会议论文信息源	221
8.3.1	《ISI Proceedings》	221
8.3.2	《世界会议》(World Meeting), Macmillan Publishing Company	221
8.3.3	中国学术会议信息源	222
8.4	学位论文信息源	222
8.4.1	网络版《ProQuest Digital Dissertations》(PQDD)	222

8.4.2	国内学位论文信息源	226
8.5	标准文献信息源	229
8.5.1	《中国国家标准库》	229
8.5.2	国际标准和国际标准化组织	229
9	MDL 信息系统的信息源	232
9.1	关于 MDL	232
9.2	Cross Fire Web	233
9.3	Cross Fire Beilstein Database	233
9.3.1	概述	233
9.3.2	检索方法和相应步骤	234
9.4	Cross Fire Gmelin Database	239
9.4.1	概述	239
9.4.2	检索方法	239
9.5	AutoNom	240
9.5.1	概述	240
9.5.2	AutoNom 的使用方法	240
9.6	《贝尔斯坦有机化学手册》和 SAN DRA	241
9.7	Beilstein Information System 网站	241
9.7.1	概述	241
9.7.2	Beilstein Abstracts on ChemWeb	242
9.7.3	联机检索数据库(Beilstein Online)	244
10	化学及相关学科数据事实信息源	245
10.1	概述	245
10.1.1	国内流行的分类	246
10.1.2	盖茨分类	246
10.2	化学及相关学科数据事实分类	247
10.3	化学及相关学科领域各类参考信息源的特点和基本功能	248
10.3.1	词典(辞典)(dictionary)	248
10.3.2	百科全书(大全)(encyclopedia)	248
10.3.3	年鉴(yearbook, annual)	249
10.3.4	手册(handbook, manual)	249
10.3.5	名录	249

10.4	选择化学及相关学科参考信息源的基本原则和方法	250
10.4.1	针对性原则	250
10.4.2	时效性原则	250
10.4.3	比较性原则	251
10.5	电子型数据事实信息源	251
10.6	化学及相关学科数据事实信息源选目	252
10.6.1	词典	253
10.6.2	百科全书	254
10.6.3	手册	255
10.6.4	机构指南、公司和产品名录	257
11	其他重要信息源选目	259
11.1	综合性信息源选目	259
11.1.1	剑桥科学文摘社及其信息源	259
11.1.2	《中国化工文摘数据库》(HGW2)	264
11.1.3	《中国化学文献数据库》	264
11.1.4	Analytical Abstracts, Royal Society of Chemistry	264
11.2	生物学科专用信息源	264
11.2.1	Biological Abstracts 和 BIOSIS Previews	264
11.2.2	CSA 所含的生物学科数据库	265
11.2.3	The Genome Database(基因组数据库)	265
11.2.4	CMS 分子化学信息源	266
11.3	材料学科专用信息源	266
11.3.1	CSA 所拥有的材料学科信息源	266
11.3.2	Solid State and Superconductivity Abstracts, 1981—Current	267
11.3.3	国内材料学科信息源	267
11.4	环境学科专用信息源	268
11.4.1	CSA 拥有的环境学科信息源	268
11.4.2	万方数据资源系统中的环境学科信息源	268
11.5	化学及相关领域企业经济信息源	268
11.5.1	DIALOG 的两个商业经济类数据库	268
11.5.2	万方资源系统商务信息子系统	269

11.5.3	几个中国经济信息方面的网站	270
11.6	索取原始文献的信息源	270
11.6.1	中国期刊网全文数据库	270
11.6.2	OCLC FirstSearch 拥有的索取原始文献 的信息源	271
11.6.3	中国高等教育文献保障系统	274
11.6.4	全国期刊联合目录网站	274
12	信息分析与组织管理	275
12.1	概述	275
12.2	信息分析、评估	277
12.2.1	信息源的分析评价及选择	277
12.2.2	检索结果宏观分析和评估方法	278
12.2.3	检索结果微观分析和评估方法	279
12.3	信息组织管理	282
12.3.1	概述	282
12.3.2	用户信息组织管理系统应具备的功能	282
12.3.3	用户信息组织管理系统的组成	283
12.4	信息组织管理系统实例——Reference Manager	285
12.4.1	Reference Manager 的主要功能	285
12.4.2	Reference Manager 的应用	286
12.5	用户信息产品	294
12.5.1	概述	294
12.5.2	文献综述	295
12.5.3	文献述评	296
12.6	信息交流	296
12.6.1	利用个人网站发布信息	296
12.6.2	在互联网上发布信息的商业软件—— Reference Web Poster	297
附录 A	Dialog 系统中与化学及相关领域有关的数据库	298
附录 B	Ei CompendexWeb 字段代码表	301
附录 C	相关网址	304
参考文献		314

绪 论

1.1 研究对象

由于文献信息源是实施检索达到检索目的的主要信息源形式,因此本书着重讨论化学及相关学科领域文献信息检索分析与管理的基本原理、手段和方法,包括检索提问分析,表述,信息源选择,制定检索策略(检索点、检索途径、检索式,选择检索手段实施检索),以及检索结果分析、评价、管理等活动应遵循的基本原理和基本方法。由于化学及相关学科领域文献信息源是实施检索的基础,因此本书将介绍和评述涉及化学及相关学科领域的主要文献信息源,并全面介绍有关文献信息源的基本知识。

此处,必须指出:文献信息检索、分析与管理所遵循的原理和方法与一般意义上的信息检索、分析与管理所遵循的原理和方法是基本一致的,即本书所述检索原理和方法可为一般信息检索、分析与管理参照。

由于信息源的电子化及针对计算机技术网络技术在信息分析与管理过程中的广泛应用,本书着力讨论电子化(数码)信息的分析与管理。

信息处于“爆炸”性增长状态,存在被信息淹没的危险,然而,知识经济对信息需求与日俱增,因此,准确、快速获取所需信息便成了人类面临的重要课题。始于 20 世纪 60 年代,在文摘-索引刊等印刷型检索工具的基础上,借助计算机技术产生了电子型检索工具——数据库。当前,随着计算机技术、数字压缩技术、通信技术(卫星通信、光纤通信)的迅速发展,联机检索、光盘检索、网络检索等已成为文献信息检索的主要手段。因此,本书将着重讨论计算机检索原理和方法。由于电子型信息源(各类数据库)是在文摘-索引刊、参考工具书等的基础上产生、发展起来的,计算机检索原理和方法与手工检索原理和方法有很多共同之点,文摘-索引刊等比数据库透明度大,直感性强,而且在一些特定场合,尚需要使用手工检索工具,因此,本书以一定篇幅讨论文摘-索引刊、参考工具书及手

工检索的原理和方法。它既是必须掌握的知识,又是学习计算机检索原理和方法的基础。基于两者的紧密联系和客观需要,采取两条原则:①以电子型信息源(各类数据库)为主,印刷型信息源(手工检索工具)为辅;②以结合、联系的方式(如将共同点抽提出来)进行论述,达到相辅相成、互相印证的目的。

信息是决策的基础,科学技术的发现和发明必须以前人、他人的知识成果为基础,这就是牛顿名言“站在巨人们的肩上”的确切含义。

21世纪是知识经济、信息时代,大学教育应强化自主学习能力、独立研究问题能力及创新精神和能力的培养,培养“学习型、创新型人才”。显然,信息素养(Information Literacy)是“学习型、创新型人才”综合素养的重要方面。那么,信息素养的含义是什么呢?众家所论略有差异,但基本之点是一致的。现引其中之一:“信息素养,它包含诸多方面:传统文化素养的延续和拓展;使受教育者达到独立自学及终身学习的水平;对信息源及信息工具的了解和运用;必须拥有各种信息技能:如需求的了解和确认,对所需文献或信息的确定和检索,对检索到的信息进行评估、组织及处理做出决定。综上所述,完整的信息素养应包括3个层面:文化素养(知识层面)、信息意识(意识层面)、信息技能(技术层面)。”^①

因此,强烈的信息意识,获取信息的知识和技能,以及分析评价信息和组织管理信息的能力,显然属信息素养范畴,是智能型人才的必备素质。本书就是服务于这一目的的。

1.2 基本体系

本书分为相互联系的四大部分:

- (1) 文献信息源的基本性状;
- (2) 检索过程的基本原理和方法;
- (3) 重要化学及相关学科文献信息源剖析;
- (4) 分析、评价检索结果和信息知识的组织管理。

这四大部分之间的内在联系:文献信息源的性状是产生检索原理和方法的基础,即只有全面深入了解文献信息源的性状,才能理解、掌握检索原理和检索方法;检索原理和检索方法是实施检索的钥匙,因此是本书的核心;然而第(3)部分所剖析、评价的化学及相关学科文献信息源是实施检索、获取所需信息的工具,是文献信息源性状的具体体现,因此,应在使用具体化学及相关学科信息源,实施检索的过程中运用、验证检索原理和检索方法,达到融会贯通之功效;分析、

^① 引自:钟伟.学与教的新坐标.中国教育报,2001-03-01