



普通高等教育“十二五”规划教材
国家精品课程主干教材

信息检索与利用

(第二版)

邓发云 编著



科学出版社

普通高等教育“十二五”规划教材
国家精品课程主干教材

信息检索与利用

(第二版)

邓发云 编著

科学出版社

北 京

内 容 简 介

本书系统地介绍了信息素养、信息源、信息检索的基本知识,以及各种常用检索工具的检索方法与技巧,提供了大量的案例、图表、思考训练题、讨论题与基于检索技巧的提示,图文并茂,便于教学与阅读。

本书突出的特点是将信息需求与信息检索结合起来,将作者长期信息检索与运用的经验与技巧贯穿其中。本书既可作为高等学校本科生、研究生的信息检索课教材与各类信息机构的岗位培训教材,也可作为工程技术人员、科研人员、教师从事信息检索时的参考书籍。

图书在版编目(CIP)数据

信息检索与利用/邓发云编著. —2版. —北京:科学出版社, 2013. 1

(普通高等教育“十二五”规划教材·国家精品课程主干教材)

ISBN 978-7-03-036313-8

I. ①信… II. ①邓… III. ①情报检索-高等学校-教材 IV. ①G252.7

中国版本图书馆CIP数据核字(2013)第001333号

责任编辑:余 江 张丽花/责任校对:包志虹

责任印制:闫 磊/封面设计:速底书装

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

文林印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2010年3月第 一 版 开本:787×1092 1/16

2013年1月第 二 版 印张:17

2013年1月第七次印刷 字数:397 000

定价:32.00 元

(如有印装质量问题,我社负责调换)

前 言

当今时代，社会分工越来越细，知识更新越来越快，工作变换越来越频繁，而这些都需要有很强的学习能力与环境适应能力，以便快速熟悉、掌握新领域。通过信息检索的学习与练习，可以有效地培养这些能力。

从目标而言，信息检索是一种交流与学习，既包括通过查询他人的或已有的历史成果与文献资料寻求答案或者激发灵感，也包括通过向他人学习、请教来释疑解惑。

从本质而言，信息检索是一种匹配和选择，是信息需求与检索结果的匹配。信息需求决定检索工具、检索方式的选择。例如：常识、大众化、新闻、娱乐信息可用搜索引擎检索；专利、标准与统计信息可从发布机构的网站检索；学术科研信息可用数据库检索；复杂、精确的需求可采用高级检索方式；模糊、简单的需求可采用简单检索方式。而检索结果的浏览、分类、排序、下载则是方便用户进行信息选择。检索工具的功能设计（检索选项或检索框）正是用来满足用户的各种需求，因此，灵活运用各种检索功能可以有效地提高检索效率。

从过程而言，信息检索是需求的明晰与调整，只有在检索的过程中才能体会与领悟到检索的意义与乐趣；只有把学习、工作、生活与科研遇到的各种问题通过信息检索的途径解决，才能不断提高信息意识与检索技能，因此，学好信息检索必须经过大量的实践训练。

本书提供了大量案例、图表、思考训练题、讨论题与基于检索技巧的提示，图文并茂，便于教学与阅读。相对于以往的大多数教材，本书增加了检索结果的分析，包括各种信息的识别、阅读、评估、运用。增加了信息需求分析、信息检索委托与求助、检索词选择等难点章节，尤其是第5章，从统一的高度来分析数据库与检索系统的检索特点与使用方法，避免了单个检索系统的大量列举。

本书是“信息检索”国家精品课程（西南交通大学）专用教材，有专门的教学网站提供背景支撑（<http://202.115.72.11>），网页中有与教材章节同步的教案、讲稿与教学设计。

本书也得到多位同志的参与和帮助，西南交通大学高凡老师编写第4章、卢世炬老师编写第6章、张蕾与王贺老师编写第8章、晓红老师编写第9章、何雪梅老师编写第10章，刘云博士、艾世尧老师参与编辑与修改，电子科技大学中山学院何海地老师编写第7章，2012年选修作者主讲本课程的五百余位学生参与素材选取、试读与修改建议，科学出版社的编辑进行了格式设计与排版，他们的付出为本书增光添彩，在此表示感谢。

在本书的内容组织、整理及写作的过程中，参阅了大量的相关文献，包括同行学者的有关论著和讲稿、各个数据库与检索系统的培训课件和使用指南、中国知网概念搜索与百度知道搜索的材料，在此谨向有关人士表示衷心感谢！同时本书也得到国家精品课程建设

资金、西南交通大学教务处和科学出版社的大力支持，一并深表感谢！

由于信息检索技术的不断发展及检索手段的不断更新，加之作者的水平有限，书中难免有不当或疏漏之处，敬请各位读者、专家及同仁批评指正。

邓发云

2012年10月

目 录

前言

第 1 章 信息检索与信息素养	1	2.2.4 联机公共检索目录及其检索	27
1.1 信息、知识、文献、情报	1	2.3 网络信息源	29
1.1.1 信息	1	2.3.1 网络的基本知识	30
1.1.2 知识	1	2.3.2 网站	32
1.1.3 文献	2	2.3.3 网络资源的分类与特点	33
1.1.4 情报	2	2.4 开放存取	34
1.1.5 信息与知识、文献、情报的关系	2	2.4.1 开放存取的特征与类型	34
1.2 信息检索概述	4	2.4.2 国内开放存取数据库及系统	35
1.2.1 信息检索的原理	4	2.4.3 国外开放存取数据库及系统	35
1.2.2 信息搜索与检索	5	2.4.4 开放存取数据库检索举例	36
1.2.3 信息检索的类型	5	2.5 联机检索系统	37
1.2.4 信息检索的方法	7	2.6 信息检索委托与求助	38
1.3 信息素养	8	2.6.1 通过文献传递、馆际互借服务方式求助	39
1.3.1 信息素养的概念	8	2.6.2 通过参考咨询方式求助	39
1.3.2 信息素养的益处	9	2.6.3 通过 E-mail 向作者索取全文	40
1.3.3 信息素养的标准	10	2.6.4 向论坛求助	40
1.4 学习信息检索的目标、意义与方法	12	2.6.5 其他求助方式	41
1.4.1 学习信息检索的目标	12	2.7 信息源使用经验与技巧	41
1.4.2 学习信息检索的意义	13	2.7.1 图书馆资源与网络资源的比较和选择	41
1.4.3 学习信息检索的方法	14	2.7.2 如何找到需要的信息资源	42
第 2 章 信息源	16	2.7.3 专业或行业网络资源的选择	44
2.1 信息源分类	16	第 3 章 信息检索技术	47
2.1.1 根据载体及制作方式的不同区分文献类型	17	3.1 信息特征、检索语言与检索途径	47
2.1.2 根据文献出版的类型和特点不同区分文献类型	18	3.1.1 信息特征	47
2.1.3 按生产加工层次划分	22	3.1.2 检索语言	48
2.1.4 各种不同类型文献的关系与对比	23	3.1.3 常用检索途径	51
2.2 图书馆	25	3.2 检索工具、数据库与检索系统	52
2.2.1 图书馆的资源与服务	25	3.2.1 检索工具	52
2.2.2 我国图书馆的类型	25	3.2.2 数据库	54
2.2.3 特色图书馆的举例	26	3.2.3 信息检索系统	56
		3.2.4 常用数据库与检索系统	60
		3.2.5 引文与引文检索系统	61

3.3 检索技术.....	62	4.2 搜索引擎的使用技巧.....	82
3.3.1 布尔逻辑检索	62	4.2.1 使用高级搜索和个性化设置	82
3.3.2 截词检索	62	4.2.2 使用类别搜索	82
3.3.3 字段检索	63	4.2.3 注意搜索返回的结果	83
3.3.4 词位置检索	64	4.2.4 注意检索策略的使用	83
3.3.5 加权检索	65	4.2.5 注意检索词的选用	84
3.3.6 检索式	65	4.2.6 其他技巧	85
3.4 检索词的选取.....	66	4.3 主要搜索引擎介绍.....	87
3.4.1 选取原则.....	66	4.3.1 Google	87
3.4.2 选取方法	66	4.3.2 百度	89
3.4.3 选取要点及技巧	69	4.3.3 雅虎中国搜索引擎	90
3.5 信息需求分析.....	71	4.3.4 搜狗搜索引擎	91
3.5.1 信息需求所涉及的通用问题	71	4.3.5 常用搜索引擎比较	91
3.5.2 用户特征所导致的信息需求 差异	71	4.4 特色搜索引擎介绍.....	92
3.5.3 不同阶段的信息需求差异.....	72	4.5 学术搜索引擎介绍.....	93
3.5.4 信息需求类型和文献类型的对 应关系	72	第5章 数据库与检索系统的使用	96
3.5.5 信息需求与信息检索的对应 关系	72	5.1 检索系统的认识与链接.....	96
3.5.6 特定信息的需求	73	5.1.1 如何认识一个数据库或者检索 系统	96
3.6 检索流程.....	73	5.1.2 检索系统的访问与连接	97
3.6.1 分析检索课题,进行信息需求 分析	73	5.1.3 数据库的合法使用	97
3.6.2 选择检索工具	74	5.2 检索系统的检索流程.....	98
3.6.3 确定检索策略	74	5.2.1 检索系统的检索界面	98
3.6.4 调整检索策略	74	5.2.2 检索系统的检索过程	100
3.6.5 获取原始文献	75	5.2.3 检索结果的处理	102
3.7 检索效果评价.....	76	5.2.4 检索系统的使用技巧	105
3.7.1 检索系统的检索效果	76	5.2.5 数据库与搜索引擎的检索特点 比较	106
3.7.2 用户的检索效果	76	5.3 检索系统的个性化服务	107
第4章 搜索引擎	79	5.3.1 创建个人账户	107
4.1 搜索引擎的原理与分类.....	79	5.3.2 个性化服务的主要功能	108
4.1.1 工作原理.....	79	5.4 常用综合检索平台	109
4.1.2 工作特点.....	80	5.4.1 综合检索系统与跨库检索	109
4.1.3 根据数据检索内容划分的搜索 引擎类型.....	80	5.4.2 中国知网	110
4.1.4 根据数据类型划分的搜索引擎 类型	81	5.4.3 万方数据资源系统	114
		5.4.4 ProQuest 平台	117
		5.4.5 EBSCOhost 平台	118
		5.4.6 OCLC FirstSearch	120
		5.4.7 Web of Knowledge	122

5.4.8 Engineering Village	126	8.1.7 专利分析	172
5.4.9 其他综合检索系统	130	8.1.8 专利检索资源推荐	173
第6章 期刊数据库检索	132	8.2 标准	175
6.1 常用中文期刊数据库	132	8.2.1 标准知识	175
6.1.1 中文科技期刊数据库(维普) ...	132	8.2.2 纸本标准资源检索	177
6.1.2 中文三大期刊全文库对比	134	8.2.3 网络标准资源检索	177
6.1.3 其他中文期刊库	134	8.2.4 标准数据库检索	179
6.2 常用外文期刊库	135	8.2.5 标准获取的其他途径	179
6.2.1 ScienceDirect (SDOL)	135	8.3 会议文献检索	179
6.2.2 SpringerLink	137	8.3.1 会议文献知识	179
6.2.3 IEL	139	8.3.2 会议文献检索途径	180
6.2.4 Wiley Online Library	142	8.4 学位论文检索	181
6.2.5 Emerald 数据库	142	8.4.1 检索特点与用途	181
6.3 其他常用期刊数据库	144	8.4.2 检索途径	181
第7章 电子图书	148	8.4.3 检索举例	182
7.1 电子图书概述	148	8.5 科技报告检索	183
7.1.1 运营形式	148	8.5.1 检索特点	184
7.1.2 与传统图书的比较	148	8.5.2 检索途径	184
7.1.3 使用方式	148	8.6 产品样本检索	184
7.2 中文电子图书	149	8.7 报纸检索	186
7.2.1 超星电子图书	149	第9章 数据与事实检索	191
7.2.2 超星“读秀”简介	151	9.1 参考工具书	191
7.2.3 方正 Apabi 电子图书	152	9.1.1 参考工具书与传统检索工具 比较	191
7.2.4 书生之家数字图书馆	152	9.1.2 字典、词典	192
7.2.5 主要中文电子图书对比	153	9.1.3 百科全书	194
7.3 外文电子图书	154	9.1.4 年鉴	195
7.3.1 MyiLibrary 电子图书	154	9.1.5 手册	196
7.3.2 Springer 电子图书	156	9.1.6 名录概述	197
7.3.3 金图国际外文原版电子图书 ...	156	9.1.7 表谱	198
7.4 其他常用电子图书简介	157	9.1.8 类书	198
第8章 特种文献检索	159	9.1.9 政书	198
8.1 专利	159	9.1.10 图录	198
8.1.1 专利知识概述	159	9.1.11 物化数据	198
8.1.2 专利检索	162	9.2 数据与事实数据库检索	199
8.1.3 国家知识产权局专利数据库 ...	164	9.2.1 万方数据资源系统	199
8.1.4 美国专利与商标局专利数 据库	167	9.2.2 中国经济信息网	199
8.1.5 欧洲专利局专利检索系统	169	9.2.3 国务院发展研究中心信息网 ...	199
8.1.6 专利检索技巧	172	9.2.4 高校财经数据库	200

9.2.5 中宏产业数据库	201	11.1.2 网站的识别	221
9.2.6 Gale Online databases	201	11.2 企业信息检索的内容与途径	222
9.2.7 BvD 财经数据库	202	11.2.1 企业名录信息检索	223
9.3 法律类信息检索	202	11.2.2 企业内部信息查找	224
9.3.1 北大法宝	202	11.2.3 企业外部信息查找	224
9.3.2 HeinOnline 法律数据库	203	11.2.4 企业评价信息查找	225
9.3.3 Lexis.com 法律专业数据库 ...	203	11.2.5 企业产品信息查找	226
9.3.4 中国法律事务应用系统	204	11.3 就业信息检索与利用	227
9.3.5 网络上的免费法律法规信息资源	204	11.3.1 以就业主管部门为主线查找信息	227
第10章 学习、考试、留学信息检索与利用	206	11.3.2 以就业去向为主线查找信息	229
10.1 学习考试类信息检索特点	206	11.3.3 就业知识信息的查找	229
10.1.1 学习考试类信息选择的关键因素——权威性	206	11.3.4 优秀就业网站推荐	230
10.1.2 专门学习考试类数据库众多 ...	206	11.3.5 就业信息的筛选	231
10.1.3 学习考试类信息选择的主要类型与途径	207	11.4 网络购物信息检索	231
10.1.4 学习考试资源的利用特点	208	11.4.1 检索途径	231
10.2 课程信息检索与利用	209	11.4.2 检索技巧	233
10.2.1 如何利用图书馆查找同类课程资源	209	第12章 学术信息检索与利用	237
10.2.2 如何查找网络上的免费课程资源	209	12.1 学术信息的识别与选择	237
10.3 考试竞赛信息的检索与利用 ...	211	12.1.1 学术信息资料的鉴别与分析	237
10.3.1 英语四、六级考试信息的检索与利用	211	12.1.2 学术信息的筛选	238
10.3.2 考研信息的检索与利用	212	12.1.3 学术信息的梳理	238
10.3.3 公务员考试信息的检索与利用	214	12.1.4 文献管理工具	238
10.3.4 竞赛信息的检索与利用	215	12.2 学术信息的阅读	239
10.4 留学信息的检索与利用	216	12.2.1 学术信息阅读的顺序选择	239
10.4.1 检索方法与技巧	216	12.2.2 学术信息阅读的精度选择	240
10.4.2 参考网站	217	12.2.3 学术信息阅读的思考与总结 ...	240
第11章 企业与商业性信息检索与利用	220	12.2.4 学术信息阅读的其他经验总结	242
11.1 企业与商业信息选择的关键因素——真实性	220	12.3 学术信息分析	242
11.1.1 商业化与生活化信息识别	220	12.3.1 信息检索与信息分析	242
		12.3.2 文献分析的目标	244
		12.3.3 文献分析的步骤	244
		12.3.4 文献分析的类型与特点	244
		12.3.5 文献信息分析成果的表达式	246
		12.4 毕业论文(设计)与学术论文	

的信息检索	246	12.5.2 引用规范	253
12.4.1 毕业论文与学术论文的要求和 特点	247	12.5.3 学术规范	253
12.4.2 毕业论文与学术论文选题	247	12.6 论文的收录、引用与发表	256
12.4.3 课题实验或设计初始阶段	249	12.6.1 收录、引用	256
12.4.4 课题实验或设计阶段	250	12.6.2 核心期刊与影响因子	256
12.4.5 论文撰写与完成阶段	250	12.6.3 国内外常见论文评价体系	256
12.4.6 各阶段检索重点对比	251	12.6.4 论文投稿与发表	257
12.5 写作与学术规范	252	本书使用方法推荐	259
12.5.1 写作规范	252	参考文献	261

第1章 信息检索与信息素养

在现代网络环境与经济全球化背景下，是否具有良好的信息素养，能否有效地获取、利用所需信息，正逐步成为衡量高素质人才自主学习能力、创新能力、创业能力的重要标准。

1.1 信息、知识、文献、情报

1.1.1 信息

在我国，“信息”一词最初来源于南唐诗人李中《暮春怀故人》一诗：“梦断美人沈信息，目穿长路倚楼台。”不过古人所说的“信息”是指“音讯、消息”。

而《现代汉语词典》中的“信息”是泛指一切音信和消息。

控制论创始人维纳认为：信息是我们用于适应外部世界，并且在使这种适应为外部世界所感知的过程中，同外部世界进行交换的内容的名称。

信息论创始人申农认为：信息是能够用来消除不确定性的东西。

日本有学者从英语词源上分析了信息（Information），认为：in 是接收到消息，formation 是整理成章。也就是说，信息是把不明确的事物弄清楚后，整理成章传递给需要的人。

英国学者阿希贝认为：信息的本性在于事物本身具有变异度。

国内学者陈原认为：信息是被一定方式排列起来的信号序列，是意义的载体。

“信息”一词的定义之所以如此多样化，主要基于以下原因：信息本身的复杂性；信息科学的新兴性，并且有诸多的学科分支；定义信息的角度不同。

在现代社会中，网上信息的发布具有很大的自由度和随意性，这必然导致信息的质量参差不齐。面对眼花缭乱的海量信息，如果处置不当，甄别不明，就会导致可怕的“信息致盲”。因此，如何合理有效地利用信息就变成一门学问。

1.1.2 知识

《现代汉语词典》中的“知识”（Knowledge）是指人们在改造世界的实践中获得的认识和经验的总和。“知识”是对某个主题确信的认识，并且这些认识拥有潜在的能力为特定目的而使用。从信息角度看，知识源于表征事物属性和事物间关系的各种信息，成于各种信息的集合或有序化。

1958年，迈克尔·波兰尼（Michael Polanyi）提出人类的知识有两类：①显性知识，即能够被人类以一定符码系统（最典型的是语言，也包括数学公式、各类图表、盲文、手势语、旗语等诸种符号形式）加以完整表述的知识；②隐性知识，其与显性知识相对，是指那种人类可以感知但难以言述的知识。

缪尔·约翰逊（Samuel Johnson）将“知识”分成两类：①人类要掌握的学科知识；②要知道在何处获得有关知识的信息。

经济合作与发展组织（OCED）将知识分为三类：①Know-What，关于事实的知识；②Know-Why，自然原理和规律方面的科学理论，形成于专门研究机构和大学等；

③Know-How,指做某些事情的技艺和能力,企业的技术情报和商业信息被归入这一类信息。

个体的知识既来源于对客观世界的观察和探索,又来源于其他个体(包括前人)的知识。因此,必须阅读科学文献,掌握有关的思想、事实、理论和方法等信息,才能在此基础上进行进一步的分析、综合和研究,并有所创新。

1.1.3 文献

《现代汉语词典》中的“文献”(Document)是泛指“有历史价值或参考价值的图书资料”。

《中华人民共和国国家标准·文献著录总则》(GB3792.1—83)为“文献”下了一个简明的定义:文献是“记录有知识的一切载体”。这就是说,文献具有两个要素:其一,知识内容。未记录有知识的空白纸张、空白磁带、磁盘、光盘等载体不能算是文献;胡乱涂写、存储的无实质知识内容的非空白纸张、磁带、磁盘、光盘等也不是文献。其二,物质载体。存在于人脑中或人们口头传播的知识,也不能算是文献。

记录知识的物质载体,既包括古代的甲骨卜辞、金石简牍、帛书卷轴、手抄文稿、书画文物、线装古籍等,也包括近、现代的印刷书刊、文书档案、电影胶卷、缩微胶片、录音带、录像带、计算机磁盘与光盘等。文献信息和载体的关系,既具有不可分割性,又具有相对独立性,也就是说,信息内容不会因为载体形式的改变而改变,不同的信息载体可以传播同一内容的信息。

文献所传递的信息是人对客观世界的反映,因而不一定完全符合客观世界表现出的信息内容,这种“歪曲”、“失真”的程度取决于人们的认识水平、立场观点、思考方法和时代因素的差异。

1.1.4 情报

《现代汉语词典》中的“情报”(Intelligence)特指战时关于敌情的报告。通常所说的情报是指运用一定的载体,传递给特定用户,用以解决科研、生产、经营中的具体问题的特定知识和信息。

情报是知识的有序化与激活。它的内核是信息,是信息的集合(知识)。杂乱无章的信息显然不是情报,再重要的信息也必须经过传递才有可能成为情报。

在经济全球化的背景下,关于技术、经济的“竞争情报”(Competitive Intelligence)受到高度重视。实际上,竞争情报的实质是组织(企业、团体乃至国家)为赢得竞争优势所需要的、具有对抗性的重要信息。

1.1.5 信息与知识、文献、情报的关系

宇宙间时时刻刻都在产生着信息,人们正是通过对不同信息的获取来认识不同事物,并由此生产新的知识。知识是经人脑思维加工而成为有序化的人类信息;文献则是被物化了的知识记录,是被人们所认知并可进行长期管理的信息;情报是人们为解决特定问题而被活化了的更为高级、更为实用的知识。情报蕴含在文献之中,但并非所有文献都是情报,而所有情报都是知识。文献又是储存和传递知识、情报和信息的介质,它们之间的逻辑关系是一种包含关系,见图1-1。

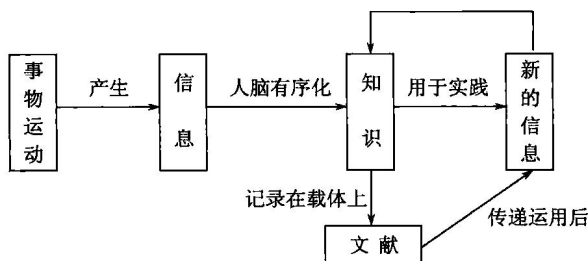


图 1-1 信息、知识与文献的关系示意图

简单地说，知识是系统化的信息；文献是静态记录的知识；情报是动态传递的知识。它们在一定条件下是可以相互转化的，它们之间的关系示意图 1-2。目前，学术界比较一致的看法是：信息>知识>情报。文献和情报的关系十分密切，而且有所交叉。

“信息”的生命过程中有两次转化：一是信息转化为知识、文献和情报；二是知识、文献和情报转化为信息。

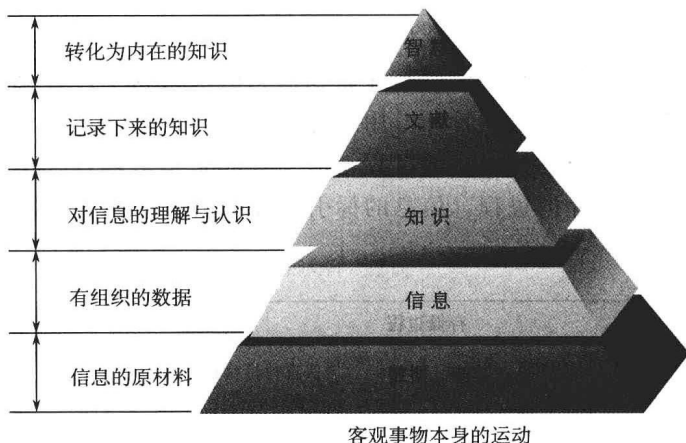


图 1-2 信息、知识、文献与智慧的关系示意图

数据是对客观事物本身运动的记录，是信息的原材料；信息是知识得以形成和传播的中介；知识是对信息的理解与认识；智慧是对知识合理的运用；所以信息不完全是知识，知识也不完全产生智慧，没有经过自我思考与消化的信息与知识无法自行凝练为个体的智慧。

尽管信息的聚合、过滤、检索诸技术可节省信息处理的时间，但信息却不应只限于被记录与管理，信息需要被思考、被挖掘与利用，才能转化为有价值的情报，转化为指导实践的智慧。

【案例 1-1】 国家统计局每个月都要公布国民经济运行数据（信息）。如果经济数据连续下滑，国家就可能采取经济刺激措施（知识）。综合各种观察，假如本月数据可能继续下降（情报），企业提前调整经营与投资策略，等待经济刺激措施出台（情报转换为决策）。

思考与讨论：如何将知识转化为智慧？如何从数据、事实提出对己和对企业有用的情报？前者需要将知识吸收与内化并与实践结合，后者必须经过长期观察与经验积累，这都需要具有一定的信息素养。

1.2 信息检索概述

信息检索就是从信息集合中找出所需信息的过程,也就是我们通常所提及的信息查寻(Information Retrieval 或 Information Search)。信息检索能力是信息素养的集中表现,提高信息素养最有效的途径则是通过学习信息检索的基本知识,进而培养自身的信息检索能力。

1.2.1 信息检索的原理

信息检索的基本原理是用户信息需求与文献信息集合的比较和选择,是两者匹配的过程。信息检索就是用户从特定的信息需求出发,对特定的信息集合采用一定的方法、技术手段,根据一定的线索与规则从中找出相关的信息。

需求集合:人们为了满足某种需求时,感到需要补充知识的集合体。

信息集合:是有关某一领域的文献或数据的集合体,它可以是数据库的全部记录,也可以是某种检索工具,还可以是某个图书馆的全部馆藏或者某个特定的信息源,它是一种公共知识结构,能弥补用户的知识结构缺陷。

匹配和选择:是一种机制,涉及两者一致性、相关度等问题,负责把需求集合和信息集合进行比较,然后根据一定的标准选出需求的信息。

广义的信息检索应包括信息的标引与存储和信息的检索两个过程见图 1-3。信息标引是指对海量的无序信息按照一定的特征,用特定的标引语言进行著录、标记和组织,使之有序化,形成用户可供检索的检索点的过程;信息存储是指对经过标引的信息进行筛选,形成检索文档和信息数据库的过程。信息的标引和存储是信息组织人员后台建立检索系统的过程。

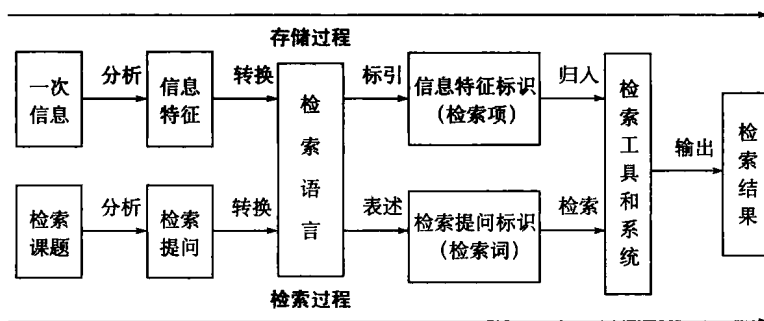


图 1-3 信息检索的一般过程

检索表达实际上是指用户将自己的需求,按照系统提供的方法和要求,将检索词用逻辑运算符联结起来,形成系统可理解和运算的查询串的过程。它主要由检索词、逻辑运算符、检索指令(检索语法)等构成。检索词是检索式的主体;而逻辑运算符和检索指令则根据具体的查询要求,从不同的角度对检索词进行检索限定。

怎样才能保证信息存得进又取得出呢?那就是存储与检索所依据的规则必须一致。也就是说,标引者与用户必须遵守相同的标引规则。这样,无论什么样的标引者,只要对同一篇文献的标引结果一致,不论由谁来检索,都能查到这篇文献。

1.2.2 信息搜索与检索

上述所说信息检索,是学术或者专业意义上的“检索”。而在现实生活中,个体一经面临各种需求,都需要搜索信息,比如找工作、找对象、购物、出门旅游、学习等,这些信息的获取可能没有严格与规范的检索系统,没有确定与明确的匹配方法,就连检索需求可能都是游弋不定的。但是,只要是检索或者查找,就应该有一些规律可循,并有一些方法可用。

其共同点是:①目标的定位,即需求的确定;②在何处找,即应该使用何种信息源,这个信息应该在哪里;③如何找,即找寻的方法与策略;④结果是否满意,即找寻的结果如何处理,如不满意,如何调整。

其不同点见表 1-1。

表 1-1 检索与搜索、查寻的区别

对比项	检 索	查找、搜索、搜寻
英文	Retrival	Search
过程和方法	有一定的策略,是系统的查找资料	随机或更随意一些
技能	需要一定的专门知识和技能	简单,任意词
用途	课题或专题	日常生活
结果	检索前通常不知道会有什么结果	通常知道结果
效率	迅速、准确	一般

【案例 1-2】 比如购房,如果个体能确定需求的价位、户型、区域、环境,在房交会、当地的房管局与房地产交易中心、媒体获取购房信息,了解网络查寻、朋友介绍、实地考察等获取购房信息的方式,知道对获取到的信息进行分析与权衡,并与自己的购买力、购买需求进行匹配,那个体的购房会相对理性与轻松,不会被促销误导或产生盲目冲动的购买行为。购房如此,找对象也是如此,所以检索信息素养可以在学习、生活与工作中不断地培养。

1.2.3 信息检索的类型

信息检索具有广泛性与多样性,根据各种具体信息检索的特点,可以将信息检索从内容、手段与检索方式等维度进行细分(见图 1-4)。

1. 按检索内容分,有数据信息检索、事实信息检索和文献信息检索

(1) 数据信息检索(Data Information Retrieval)是将经过选择、整理、鉴定的数值数据存入数据库中,根据需要查出可回答某一问题的数据的检索。其检索对象既包括物质的各种参数、电话号码、银行账号、观测数据、统计数据等数字数据,也包括图表、图谱、市场行情、化学分子式、物质的各种特性等非数字数据。数据检索是一种确定性检索,信息用户检索到的各种数据,是经过专家测试、评价、筛选过的,可直接用来进行定量分析。例如,检索“2011 年中国国民生产总值是多少?”,“中国移动的服务电话是多少?”,等等。

(2) 事实信息检索(Fact Information Retrieval)是将存储于数据库中的有关某一事件发生的时间、地点、经过等情况查找出来的检索。其检索对象既包括事实、概念、思想、知识

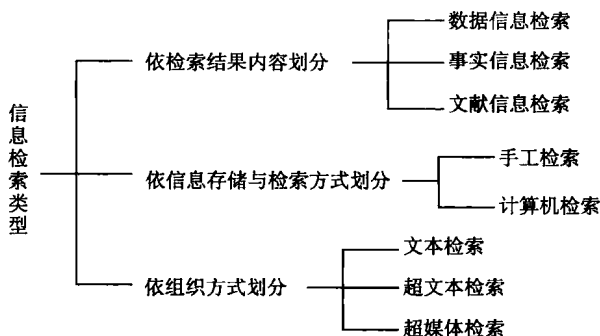


图 1-4 信息检索的类型

等非数值信息，也包括一些数据信息，但需要针对查询要求，由检索系统进行分析、推理后，再输出最终结果。例如，检索“联想集团的创始人是谁？它在哪个交易所上市？”。

(3) 文献信息检索 (Document Information Retrieval) 是将存储于数据库中的关于某一主题文献的线索查找出来的检索。检索结果往往是一些可提供研究课题使用的参考文献的线索或全文。文献检索是信息检索的核心部分。根据检索内容不同文献检索又可分为书目检索和全文检索。例如，“关于地震与海啸有些什么关联？”这就需要检索主体根据课题要求，按照一定的检索标识（如主题词、分类号等），从所收藏的文献中查出所需要的文献。

提示：在数据检索和事实检索中，用户需要获得的是某一事物或某一数据的具体答案，是一种确定性检索，一般利用参考工具书；如果检索的事物与数据是一些大众化、公开性或者常识类信息则可通过搜索引擎直接查寻。文献检索通常是检索所需要信息的线索，需要对检索结果进行进一步分析与加工，一般使用检索刊物、书目数据库或全文数据库。

2. 按组织方式分，有文本检索、超文本检索和超媒体检索

(1) 文本检索 (Full Text Retrieval) 是将存储在数据库中的整本书、整篇文章中的任意内容信息查找出来的检索。可以根据需要获得全文中的有关章、节、段、句、词等的信息，也可以进行各种统计和分析。

(2) 超文本检索 (Hyper Text Retrieval) 是对每个节点中所存的信息以及信息链构成的网络中信息的检索。强调中心节点之间的语义联接结构，靠系统提供的工具进行图示穿行和节点展示，提供浏览式查询，可进行跨库检索。

(3) 超媒体检索 (Hyper Media Retrieval) 是对存储的文本、图像、声音等多种媒体信息的检索。它是多维存储结构，有向的链接，可提供浏览式查询和跨库检索。

3. 按检索手段分，有手工检索和计算机检索

(1) 手工检索 (Hand Retrieval) 是检索人直接用手、眼、脑组织查找印刷型文献的检索。这种检索方式具有直观、灵活，无需各种设备和上机费用的优点。

(2) 计算机检索 (Machine Retrieval) 是通过机器对已数字化的信息，按照设计好的程序进行查找和输出的过程。按处理方式分又有脱机检索和联机检索；按存储方式分，有光盘检索和网络检索。计算机检索的优势在于，可大大提高检索效率，拓宽检索领域。

提示：按 Ctrl+F 键，出现查找搜索框，键入所要查找的关键词，可在当前桌面中查找。可比肉眼直接查找快多了。

1.2.4 信息检索的方法

信息检索的方法有多种，分别适用于不同的检索目的和检索要求。归纳起来，常用的信息检索方法有常规检索法、回溯检索法、循环检索法。

1. 常规检索法

常规检索法又称常用检索法、工具检索法。它是以主题、分类、作者等为检索点，利用检索工具获得信息资源的方法。根据检索方式，常规检索法又分为直接检索法和间接检索法；根据检索需求，常规检索法又分为顺查法、倒查法和抽查法。

(1) 直接检索法。这是指直接利用检索工具进行信息检索的方法。例如，利用字典、词典、手册、年鉴、图录、百科全书等进行检索。这种方法多用于查找一些内容概念较稳定、较成熟、有定论可依的指示性问题的答案。

(2) 间接检索法。这是利用检索工具间接检索信息资源的方法，具体包括三种（见表 1-2）。

表 1-2 三种间接检索方法对比

类型	定 义	适用范围	特 点
顺查法	根据检索课题的起始年代，利用选定的检索工具按照由远及近、由过去到现在顺时序逐年查找，直至满足课题要求为止	普查一定时间的全部文献，查全率较高，并能掌握课题的来龙去脉，了解其研究历史、研究现状和发展趋势	方法费力、费时，工作量大，多在缺少评述文献时采取此法。因此可用于事实性检索
倒查法	与顺查法相反	多用于新课题、新观点、新理论、新技术的检索，检索的重点在近期信息上，只需查到基本满足需要时为止	查到的信息新颖，节省检索时间。但查全率不高，容易产生漏检的现象
抽查法	针对某学科的发展重点和发展阶段，拟出一定时间范围，进行逐年检索的一种方法	根据检索需求，针对所属学科处于发展兴旺时期的若干年进行文献查找	检索效率较高，但漏检的可能性大，检索人员必须熟悉学科的发展特点

2. 回溯检索法

回溯检索法又称追溯法、引文法、引证法，是一种跟踪查找的方法。这种检索方法不是利用确定的检索工具，而是利用已知文献的某种指引（如文献附的参考文献、有关注释、辅助索引、附录等）追踪查找文献。用追溯法检索文献，最好利用与研究课题相关的专著与综述。在检索工具不全或文献线索很少的情况下，可采用此法。

常见的追溯方式有：文章→参考文献→更多文章；作者→团体→更多作者→文章；链接→网站→更多链接；专利→发明人→论文；专利→申请人→专利等。

另外，还有一些专门用于追溯法的检索工具，即引文索引。这类检索工具比较著名的有美国的《科学引文索引》和《中国社会科学引文索引》。由于追溯法的有效性，目前一些非引文检索工具也采用追溯法的思想，将众多的文献关联起来。比如，在中国知网（CNKI）的各个数据库检索结果中，就有参考文献、引证文献、相似文献、读者推荐文献等。