

现代信息检索

过仕明 主编

黑龙江教育出版社

图书在版编目(CIP)数据

信息检索教程 / 过仕明主编. — 哈尔滨: 黑龙江教育出版社, 2004.6

ISBN 7-5316-4244-1

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

中国版本图书馆 CIP 数据核字(2004)第 063242 号

现代信息检索

XIANDAI XINXI JIANSUO

过仕明 主编

责任编辑: 贾海涛

责任校对: 王 军

封面设计: 傅 旭

黑龙江教育出版社出版(哈尔滨市南岗区花园街 158 号)

东北林业大学印刷厂印刷·黑龙江教育出版社发行

开本 787 × 1092 毫米 1/16·印张 22.5·字数 500 千

2004 年 8 月第 1 版·2004 年 8 月第 1 次印刷

印数: 1 - 2000

ISBN 7-5316-4244-1/G·3199 定价: 38.00 元

如有质量问题, 请与印刷厂联系退换。

序 言

伴随着 21 世纪的到来,人类社会已由工业化时代进入信息化时代,信息资源、能源和物质一起成为现代信息社会的三大支柱。由于计算机的广泛应用,人类能够高效率地开发并利用信息,以计算机技术、通信网络技术为核心的信息技术已成为推动社会发展的强大动力。因此信息的数字化、数字化信息的网络化,网络环境下海量信息的有序组织、存储和检索,即信息检索将是我们今后相当长一段时间内必须关注的前沿课题。

随着信息技术的发展,尤其是信息网络的大规模建设和实际应用,用户的各项需求更多地转向依靠电子书刊、电子数据库、E-mail 等电子信息资源网络而得以满足,信息检索也经历了从手工检索到计算机检索再到今天的因特网检索的发展过程。而且随着网络信息环境的出现,使得原有信息检索研究的对象和范围随之改变并扩大,研究信息资源的群体也改变了由原有的以图书情报领域的专家学者为主的框架,众多的信息公司、咨询公司等加入到研究开发信息检索系统的行列。网络使计算机信息检索技术进入了一个崭新阶段,而网络信息检索又使得网上信息源利用率提高,信息组织更为有序和高效。

正是由于信息检索范围及检索工具、方式、处理技术的不断发展,因此有必要对现代信息检索技术和发展进行总结,《现代信息检索》一书正是为了推广和普及检索技术而编著的。

本书较为系统地介绍了信息检索的基本内容,是一本介绍信息检索原理和方法及中外各学科最常用、最重要的检索工具(包括数据库、搜索引擎等)的教学参考书。主要介绍包括信息检索的原理、方法及意义,检索语言、检索工具的编制和评价,检索工具的使用,手工检索、光盘检索、联机检索和网络检索。

在编写中注重了教材的通用性、可行性和易操作性,力求通俗易懂,简便实用,着重介绍信息资料查找的基本知识和入门的必备知识,旨在培养学生的信息意识和信息查找能力,使其能够及时、准确、有效地利用各种信息,增强在激烈的竞争中生存和发展的能力。

本书由过仕明同志主编,各章节的编写具体分工如下:

过仕明编写第一至第四章,张明华编写第六、七章,孔翔雁编写第十二、十三章,王磊编写第五、八章,杨翠屏编写第九、十章,杨晓秋编写第十四章和第十一章的 1~4 节,李丹编写第十五章和第十一章的 5~7 节。

全书由过仕明同志统稿。

本书的编写参阅和引用了许多作者的研究成果和来自互联网的内容,我们都以参考

文献和引用文献的形式列于书末，他们的研究成果为本书提供了丰富的素材，在此表示由衷的感谢，同时也感谢徐广鑫同志对本书提供的支持。

由于作者水平有限，本书必然会存在不少缺陷，必然要不断修改、不断完善，恳请读者批评指正。

编者
2004年7月

目 录

第一章 信息检索概论	1
1.1 信息概论	1
1.1.1 信息概念与特性	1
1.1.2 信息的几个相关概念	4
1.2 信息检索概念与类型	11
1.2.1 信息检索的概念	11
1.2.2 信息检索的类型	12
1.3 信息检索原理	14
1.4 信息检索的方法与途径	17
1.4.1 信息检索方法	17
1.4.2 信息检索途径	19
1.5 信息检索步骤	22
第二章 信息检索技术与效率评价	27
2.1 信息检索发展历程	27
2.2 信息检索技术方法	31
2.2.1 手工信息检索技术方法	31
2.2.2 计算机信息检索技术方法	43
2.3 检索效率评价	45
第三章 信息检索语言	49
3.1 信息检索语言的概况	49
3.1.1 检索语言的概述	49
3.1.2 检索语言的分类	51
3.2 分类语言	51
3.2.1 分类语言概述	51
3.2.2 分类语言的类型	51
3.2.3 我国分类法的简介	53
3.3 主题检索语言	58
3.3.1 主题语言的概述	58
3.3.2 主题词的类型	59
3.4 分类语言与主题语言	69

3.4.1	分类语言与主题语言的差别	69
3.4.2	主题语言与分类语言的联系	71
第四章	工具书及其检索	72
4.1	工具书的特点及功用	72
4.1.1	工具书的定义	72
4.1.2	工具书的特点	72
4.1.3	工具书的类型	73
4.2	检索工具	73
4.2.1	检索工具的定义、职能及特征	74
4.2.2	检索工具的类型	74
4.3	参考性工具书	79
4.3.1	参考性工具书的概述	79
4.3.2	参考性工具书检索事例	88
4.4	工具书的基本结构及排检方法	89
4.4.1	工具书的基本结构	89
4.4.2	中文工具书的排检方法	89
第五章	外文科技期刊文献检索	96
5.1	美国《工程索引》及其检索	96
5.1.1	EI 的概况	96
5.1.2	EI 的编排结构	97
5.1.3	EI 的著录格式	97
5.1.4	EI 的检索工具	99
5.1.5	EI 检索的途径	100
5.2	英国《科学文摘》及其检索	100
5.2.1	SA 的概况	100
5.2.2	SA 的编排结构	101
5.2.3	SA 的著录格式	101
5.2.4	文摘的辅助索引	102
5.3	美国的《化学文章》及其检索	103
5.3.1	CA 的概况	103
5.3.2	CA 的编排结构	104
5.3.3	文摘著录格式	105
5.3.4	辅助索引体系	105
5.4	美国的《科学引文索引》及其检索	107
5.4.1	基本概念	107
5.4.2	《科学引文索引》的概况	108
5.4.3	SCI 的结构及著录格式	109
5.4.4	SCI 的检索途径	110

第六章 专利文献及其检索	112
6.1 专利及专利制度	112
6.1.1 专利的基本概念	112
6.1.2 专利的类型	112
6.1.3 专利权的基本特征	113
6.1.4 授予专利权的条件	114
6.1.5 专利制度	115
6.2 专利文献及其检索	116
6.2.1 专利文献	116
6.2.2 专利文献的特点	116
6.2.3 专利文献的作用	117
6.2.4 专利说明书	118
6.2.5 常用检索工具	118
6.3 《国际专利分类表》介绍	119
6.3.1 国际专利分类法的作用	120
6.3.2 《IPC》分类原则	120
6.3.3 《IPC》的分类体系结构	120
6.3.4 《IPC 的关键词索引》	121
6.4 中国专利文献及其检索	121
6.5 美国专利文献及其检索	124
6.5.1 美国专利概述	124
6.5.2 美国专利的印刷型检索工具	124
第七章 特种文献及其检索	127
7.1 学位论文及其检索	127
7.1.1 学位论文概述	127
7.1.2 学位论文写作的基本格式	128
7.1.3 学位论文写作程序和基本要求	130
7.1.4 学位论文选题的基本原则	132
7.1.5 常用学位论文检索工具	133
7.2 会议文献及其检索	137
7.2.1 学术会议与会议文献	137
7.2.2 会议文献的类型	138
7.2.3 会议文献的检索工具	139
7.3 科技报告及其检索	140
7.3.1 科技报告概述	140
7.3.2 科技报告的类型和特点	141
7.3.3 科技报告的产生过程	143
7.3.4 美国政府四大科技报告及其检索	144

7.4 标准文献及其检索	146
7.4.1 标准文献概述	146
7.4.2 标准文献的类型和作用	147
7.4.3 国内标准及检索	148
7.4.4 国际标准文献及检索	150
7.4.5 网上标准文献检索	151
第八章 计算机信息检索	153
8.1 计算机信息检索概述	153
8.1.1 计算机信息检索系统的发展历史	153
8.1.2 计算机信息检索系统的发展方向	155
8.1.3 我国计算机信息检索系统的发展	156
8.2 计算机信息检索系统基础	157
8.2.1 计算机信息检索系统的组成	157
8.2.2 计算机信息检索原理	159
8.2.3 计算机信息检索与手工检索的比较	159
8.2.4 计算机信息检索系统的服务方式	160
8.3 机读数据库介绍	160
8.3.1 信息存储与数据库结构	160
8.4 计算机检索技术	165
8.5 计算机检索的程序与策略	170
第九章 国际联机检索	174
9.1 国际联机检索概述	174
9.1.1 国际联机检索系统发展历史	174
9.1.2 国际联机检索系统的组成	175
9.1.3 国际联机检索系统的特点	176
9.1.4 国际联机检索提供的服务	176
9.1.5 国际联机数据库的类型和结构	177
9.1.6 国际联机检索操作基本过程	177
9.1.7 国际联机检索的检索方法	178
9.2 国际著名联机检索系统简介	178
9.3 DIALOG 联机检索系统	181
9.3.1 DIALOG 提供的专门搜索技术	182
9.3.2 DIALOG Web 用法	189
9.4 STN 联机检索系统	201
9.5 Questel - ORBIT 系统	204
9.6 ESA - IRS 系统	212
9.7 国内联机信息检索系统	214
9.7.1 中国科技信息研究所联机信息检索系统	214

9.7.2	北京文献服务处联机检索系统 (BDS)	214
9.7.3	国家图书资料信息光盘网络 (National Publications CD - ROM Net)	214
第十章	光盘信息检索	215
10.1	光盘概述	215
10.1.1	光盘及光盘的结构	215
10.1.2	光盘的类型	216
10.1.3	光盘的特点	217
10.2	光盘检索系统	218
10.2.1	光盘检索系统的发展阶段	218
10.2.2	光盘检索的过程和常用方式	219
10.2.3	光盘检索系统提供的服务	220
10.3	光盘数据库	220
10.3.1	光盘数据库的种类	220
10.3.2	光盘数据库的特点	221
10.3.3	光盘数据库检索与国际联机检索的比较	221
10.4	数据库的发展	222
10.4.1	数据库业	222
10.4.2	美国的数据库业	223
10.4.3	西欧的数据库业	225
10.4.4	中国的数据库业	226
10.4.5	数据库的新发展	227
第十一章	国内外主要数据库选介	229
11.1	《中国学术期刊》全文数据库检索	229
11.1.1	概况	229
11.1.2	检索界面	230
11.1.3	检索方法	232
11.1.4	检索结果的处理	233
11.2	《复印报刊资料》光盘数据库	235
11.2.1	概况	235
11.2.2	《人大复印报刊》光盘资料数据库使用说明	235
11.3	《中文科技期刊数据库》	238
11.3.1	概况	238
11.3.2	《中文科技期刊数据库》使用说明	238
11.4	其他国内外光盘数据库简介	240
11.4.1	《中文期刊数据库》	240
11.4.2	《人民日报》光盘数据库	241
11.4.3	《参考消息四十年光盘》数据库	242

11.5	德国 Springer 数据库	243
11.5.1	Springer 的发展	243
11.5.2	使用途径	244
11.5.3	检索规则	248
11.5.4	其他功能	248
11.5.5	检索实例	249
11.6	EBSCO 数据库	249
11.6.1	EBSCO 包含的主要数据库	249
11.6.2	数据库基本功能及操作	251
11.6.3	检索方法	252
11.6.4	检索结果处理	256
11.6.5	EBSCO 常用算符	257
11.6.6	检索实例	257
11.7	UnCover 数据库	257
11.7.1	UnCover 的主要服务项目	258
11.7.2	检索方法	258
11.7.3	显示 CALIS 馆藏信息	259
第十二章	Internet 基础知识	260
12.1	Internet 简述	260
12.1.1	Internet 的历史与发展	260
12.1.2	Internet 在中国的发展	261
12.2	网络基础	263
12.2.1	网络的组成	263
12.3	INTERNET 应用	268
12.3.1	World Wide Web	268
12.3.2	电子邮件	270
12.3.3	Internet 服务	273
12.3.4	电子公告板和网络新闻组	275
12.3.5	电子商务	276
第十三章	网络信息资源	279
13.1	网络信息资源概述	279
13.1.1	网络信息资源的类型	279
13.1.2	网络信息资源的特点	280
13.1.3	网络信息资源存在的问题	281
13.2	电子期刊	282
13.2.1	电子期刊发展概况	282
13.2.2	电子期刊的类型	283
13.2.3	电子期刊的优点	284

13.2.4	电子期刊存在的问题	285
13.2.5	国内外主要电子期刊选介	286
13.3	电子报纸	289
13.3.1	电子报纸及其发展	289
13.3.2	电子报纸的特点	290
13.3.3	电子报纸存在的问题	291
13.3.4	电子报纸分类	291
13.3.5	电子报纸选介	292
13.4	电子图书	294
13.4.1	电子图书的特点	295
13.4.2	国外主要的电子图书	295
13.4.3	国内主要的电子图书	296
13.5	数字图书馆	297
13.5.1	数字图书馆的发展	298
13.5.2	数字图书馆的特点	299
13.5.3	国外典型的数字图书馆项目	300
13.5.4	国内典型的数字图书馆	301
第十四章	网络信息资源检索工具	303
14.1	网络信息检索工具及其类型	303
14.1.1	网络信息检索工具	303
14.1.2	网络信息检索工具的类型	303
14.2	网络信息检索工具的构成及原理	307
14.2.1	网络信息检索工具的构成	307
14.2.2	网络信息检索工具的原理	308
14.2.3	运用网络信息检索工具应注意的一些问题	309
14.3	搜索引擎	310
14.3.1	搜索引擎的起源和含义	310
14.3.2	搜索引擎的分类	310
14.3.3	搜索引擎的检索策略	311
14.3.4	使用搜索引擎的方法与技巧	311
14.3.5	搜索引擎的劣势	314
14.3.6	如何对待检索结果	314
14.3.7	正确对待搜索引擎	315
14.3.8	搜索引擎的未来发展趋势	315
第十五章	搜索引擎的使用	317
15.1	中文综合性搜索引擎	317
15.1.1	百度搜索引擎 (http://www.baidu.com/)	320
15.1.2	搜狐搜索引擎 (http://www.sohu.com/)	320

15.1.3	新浪搜索引擎 (http://ha.sina.com.cn/)	323
15.1.4	其他搜索引擎	326
15.2	英文综合性搜索引擎	328
15.2.1	Google (http://www.google.com/)	328
15.2.2	AltaVista (http://www.altavista.com/)	333
15.2.3	Yahoo! (http://www.yahoo.com/)	335
15.2.4	Excite (http://www.excite.com/)	337
15.2.5	其他英文网络检索工具	339
15.3	图像、影像及相关信息的网络检索工具	341
15.4	特殊搜索引擎	348
参考文献		350

第一章 信息检索概论

1.1 信息概论

1.1.1 信息概念与特性

1. 信息概念

信息的定义众说纷纭，莫衷一是。人们从不同学科、不同角度或用不同方法对信息下了许多定义：

(1) 信息 《辞海》1989年版解释为：音讯，消息。早在一千多年前，唐朝诗人李中在《碧云集·暮春怀故人》一诗中就留下了“梦断美人沉信息，目穿长路倚楼台”。通信系统传输和处理的对象，泛指消息和信号的具体内容和意义。通常须通过处理和分析来提取。信息和物质、能量被称为构成系统的三大要素。信息的量值与信息的随机性有关，如在接收端无法预估消息或信号中所蕴含的内容或意义，即预估的可能性越小，信息量就越大。

(2) 信息 信息论的创始人申农认为：“信息是能够用来消除不确定性的东西”。它是通过多种形式，包括数据（字母、符号和数字）代码、图形、报表、指令等等反映出来的。这是从通信理论出发来研究信息传递与度量问题。

(3) 信息 控制论创始人维纳说：“信息是人们在适应外部世界并且使这种反应作用于外部世界的过程中同外部世界进行交换的内容的名称。”

(4) 信息 美国《未来学家》杂志所刊登过的专家们撰写的文章指出：“它是一切容易获得的和不易获得的，有时可供人们参考的事实和思想的总和。”

(5) 信息 我国学者钟义信指出：“信息是事物存在的方式或通信的状态，以及这种方式、状态的直接或间接的表述。”

(6) 信息 近年来人们一般认为：信息是世界上一切事物运动的状态、特征及其反应。它与事物同在，与事物共存，存在于整个自然界与人类社会。

信息定义的多样化，主要源于三方面的原因：第一，信息本身的复杂性。它是一个多元化、多层次、多功能的复杂综合体。第二，信息科学尚不够完善成熟。它的内涵与外延不甚确切，一些重要概念正处于多定义并存阶段。第三，实际需要的不同。人们出

于不同的研究目的,从不同角度出发,对信息作不同的理解与解释。一般对信息概念通俗的解释是:“信息是物质的一种普遍属性,是物质存在的方式和运动的规律与特点”。

2. 信息的特性

信息,按其性质分为自然信息(生物信息,天体信息)和社会信息;按其载体可分为空气信息、电子信息、光纤信息和文献信息等。不论哪种类型的信息,都具有共同的特性。

(1) 普遍性

信息不仅普遍存在于自然界、人类社会以及人类思维活动之中,而且还是客观世界中一切物质的属性,它反映的是物质之间的普遍联系,这种联系是事物运动的状态和方式,只要有事物存在,只要有事物的运动,就会有其运动的状态和方式,就存在着信息。无论在自然界、人类社会还是在人类思维领域,绝对的“真空”是不存在的,绝对不运动的事物也是没有的。因此,信息是普遍存在着的。普遍性是信息的最重要本质属性和特征之一。信息与物质、能量一起构成了客观世界的三大要素。

(2) 客观性

信息不是物质,也不是能量,是客观事物运动、变化特征与规律的客观描述,信息统一于物质世界,其根源是物质世界,它与整个世界共存。

(3) 流动性与传递性

信息在事物之间的相互联系必定在信息的流动中发生,因此,它具有流动性。信息的传递性表现在人与人之间的消息交换,人与自动机、自动机与自动机之间的信息交换,动物界和植物界的信号交换,同时,人类进化过程中的细胞选择、遗传也被看做是信息的传递与交换。

(4) 无限性

信息与本源物质以及物质的运动形式一样,是无限的,取之不尽,用之不竭。物质是无限可分的,物质的层次也是无限的。某一具体的物质通过其内部结构及其与外界的联系所带有的信息是一定的,又是无限的。在人类生存和社会活动的一切领域,随时都在交流、产生和处理信息。随着时间的推移,人类所掌握和涉及到的信息的内容和范围在扩大,信息量也在无止境的增加。物质的边界是无限的,对物质世界的认识也是无限的,人的创造能力也是无限的。信息的多样性、衍生性是信息的无限性的重要特征。

(5) 可度量性

信息是可以度量的,信息的度量,犹如物质和能量的度量一样,关键在于对度量原理的认识以及在此基础上选择适当的度量标准。最早给信息以定量描述的是申农。他确定信息量的单位为比特(Bit)。1比特的信息量,在变异度为2的最简单情况下,就是能消除非此即彼的不确定性的需要的信息量。与熵是系统的无序程度的度量相反,信息量是系统的有序程度的度量,故表现为负熵。它的公式为: $H(x) = -\sum P(X_i) \log_2 P(X_i)$,式中 X_i 代表 n 个状态中的第 i 个状态, $P(X_i)$ 代表出现第 i 个状态的概率, $H(X)$ 代表用以消除系统不确定性所需的信息量,即以比特为单位的负熵。

(6) 依附性

载体是信息存在的必要条件。信息本身是看不见、摸不到的,它必须依附于一定的物质形式(如声波、电磁场波、纸张、光学材料、磁性材料等)之上,不可能脱离物质而单独存在。

上述信息的普遍性、客观性、流动性、传递性、无限性、可度量性、依附性,是信息最基本的特性。

为了更好地开发和利用信息资源,我们还应进一步了解信息在上述基本特性基础上衍生而成的,具有一定实际意义的相关重要性质。

(1) 多样性与综合性

信息在不同的领域具有多种不同的特性或表现形式,如客观事物中的各种自然属性;人工设备的技术特征;人类社会的各种社会特征;人脑中反映客观事物认识的思想、知识;人类交流信息过程中的声音、文字、图像以及用各种编码形式记录下来的数据、新闻、情报、消息等。各种形式的信息又常常以综合的方式表现事物的特征,所谓“多媒体”正是信息多样性和综合性的集中表现。

(2) 相对性与有效性

从信息作为事物相互联系的反映角度看,信息源不确定的程度或者信息源被接受的信息量多少,均与信宿的状态有关。这一特征在人作为信宿接受信息的过程中表现尤为突出。同一信息相对于具有不同认知水平的人所产生的作用和有效性也不相同。

(3) 积累性与价值性

信息通过人脑思维或人工技术设备的综合、处理和加工,不断积累丰富,提高其质量和利用价值。信息的质量与价值,实际上是对客观事物属性反映的深度和真实程度的认识。虽然信息是人类的一种重要资源,但信息只有被利用才会产生价值,否则,其价值或随时间的流逝而减少,或成为“信息垃圾”。

(4) 共享性

共享性是信息区别于物质和能量的主要特性。信息的共享性主要表现在同一内容的信息可以在同一时间由两个或两个以上的使用者使用。信息的共享性是建立在信息的普遍性基础上的。信息共享性实现的条件是信息对物质依附性的相对性。信息资源利用范围的扩大不仅表现在信息内容的深化和扩大,而且还表现在固定信息内容的共享范围的实现和扩大。不同的信息,共享的范围是不同的。

(5) 异步性

人们对于事物运动状态和方式的认识,即信息,严格地说来,总是产生、传达在事实之后的。先有了事实,而后才有认识,才可能有信息,信息再快,也有滞后性。即使是大脑,它在获取一定信息的基础上,经过思维,产生新的思想,发出指令并最终表现在行动上也需要一个过程,思想和指令是在思维之后,而关于行动的信息又在行动之后。但是,信息这种意义下的滞后性,并不意味着人类不能产生和利用超前于事实发生的信息。预测性信息来源于大量滞后信息的积累和人的主观能动性,它的有用性建立在对事物发生、发展的运动规律的认识和探索的基础上,它的准确性受到所积累的信息的质和量、对信息的处理分析和利用能力、预测方法、即将发生的事物所处的环境条件变化等多方面的影响,需要实际发生和发展的事物来验证。验证的结果、验证的方法等又

将为预测其他同类事物提供信息的积累。信息的滞后性是绝对的,信息的超前性是相对的。信息的滞后性和超前性构成了信息同时的异步性。

(6) 不完全性

任何关于客观事物的运动状态和方式的认识和知识都不可能包揽无余,实际上也不必要。这是由信息的无限性所决定的。由于人的认识能力有限、信息的绝对滞后性所限、搜集利用信息的成本——效益原则等决定了在任何一个给定的历史阶段,人类关于自然界和社会本身的认识都是有限的,人的认识的无限性是相对于历史的长期发展而言的。有限的认识和无限的信息相矛盾的结果,就是人类关于事物运动状态和方式的认识的不完全性。

(7) 污染性

信息用户所面临的信息“污染”是指,在包罗万象的信息海洋中寻找所需信息时,一方面很难找到正确的信息,一方面又被质量差或错误的信息所环绕。

1.1.2 信息的几个相关概念

信息、知识、情报、文献这四个概念在许多书或文章中经常会遇到,它们之间既有密切的联系又有本质的区别。为了更好地掌握信息检索的知识,下面将它们分别加以介绍。信息的概念和特征前面已经详细介绍,这里不再重复。

1. 知识

知识是人类对客观事物存在和运动规律的认识。

知识,《辞海》1989年版解释为:人类认识的成果或结晶。包括经验知识和理论知识。经验知识是知识的高级形态。人类的知识是在后天社会实践中形成的,是在社会实践中积累起来的经验,是对现实的反映。辩证唯物主义把社会实践作为一切知识的基础和检验知识的标准。知识借助于一定的语言形式,物化成为某种劳动产品的形式,交流和传递给下一代,成为人类共同的精神财富。知识随社会实践、科学技术的发展而发展。知识一般可以分成三大类:自然科学知识、社会科学知识和思维科学知识。哲学知识则是关于自然、社会和思维知识的概括和总结。

知识,也可这样解释为:知识是人类在改造客观世界实践中所获得的认识和经验的总和。人们认识客观事物的过程就是人脑对外界事物传来的信息加工的过程。而认识飞跃的结果即为知识。也就是说信息被有选择地吸收到人们的思维系统,经过人们大脑的储存、甄别、加工、处理、转换等形式而形成知识。

知识是信息的一部分,信息的概念要大于知识。人们通过信息感知世界、认识世界和改造世界,并且将获得的信息转变成知识作为认识和改造世界的武器。我们可以说知识是一种信息,而不能说信息就是知识。知识与信息可以相互转化,使用者把信息转化为知识,再把知识转化为智慧,是一种动态过程,是一种开拓过程。反过来,智慧又会转化为新知识,新知识又会转化为新信息,人们通过一定的社会传递过程或媒体把信息传递给使用者。

2. 文献

文献,《辞海》1989年版解释为:原指典籍与宿贤。《论语·八佾》:“夏礼吾能言之,杞不足徵也;殷礼吾能言之,宋不足徵也;文献不足故也。足,则吾能徵之矣。”朱熹注:“文,典籍也;献,贤也。”后专指具有历史价值的图书文物资料,如:历史文献。亦指与某一学科有关的重要图书资料,如医学文献。今又为记录有关知识的一切载体的统称。即用文字、图像、符号、声频、视频等手段将人类知识记录在各种载体(如:纸张、胶片、计算机磁盘、光盘等)上。

文献,国际标准化组织《文献情报术语国际标准》(ISO/DIS5217)的解释为:“在存储、检索、利用或传递记录信息的过程中,可作为一个单元处理的,在载体内,载体上或依附载体而存储有信息或数据的载体。”

文献,《中华人民共和国国家标准·文献著录总则》(GB3792·1-83)定义为:“文献是记录有知识的一切载体”。以上三种说法的基本含义是相同的,都强调了文献的三个基本属性,即文献的知识性、记录性和物质性。也就是说,文献既不是知识本身,也不是记录知识的物质载体本身,而是两者的结合,即:文献=知识+载体。知识是文献的内容,载体是文献的外在形式,文献是知识内容与载体形式的统一体。因而,文献与信息、知识、情报概念,从根本上来讲属于不同种类的事物。

从上述定义也很容易看出,文献所包含的范围十分广泛。不仅古代的甲骨文、碑刻、竹简、帛书是文献;图书、报纸、期刊是文献;现今的机读资料、缩微制品、电子出版物等也是文献。

3. 情报

情报与信息的英文单词尽管同是“Information”,但二者有着本质的区别,下面将分别介绍二者的异同。

一致性表现在:情报一词泛指一切最新的情况、消息的报道、交流,其本义来自于军事上的“敌情报告”,是军事行动的重要依据之一,带有军事行为,具有保密性的含义。因此,情报存在于人类社会的各项活动之中,是人类社会的一种普遍现象。从这些意义上讲,“情报”与信息概念中的“社会信息”的含义是相通的,可以说具有一致性。

鉴于在人类社会“情报”与“信息”两词的相通性,1992年国家科委决定将科技情报司改名为科技信息司,并建议全国科技系统不再使用情报一词,而改用信息二字,使之更具有时代的特点。过去所说的情报检索改为今天所讲的信息检索,使检索的内容更加广泛。但这种作法尽管有利于统一认识,但学术界大多持谨慎态度,因为“信息”和经过组织的知识性的“情报”毕竟不是同一层次的事物。

不同点表现在:从另一个角度来看,情报作为科学交流的对象是指一切传递中的能够改变认知主体原有知识结构的知识,它才是情报概念的主体构成部分。我国著名科学家钱学森指出:“情报就是为了解决一个特定的问题所需要的知识,要注意它的及时性和针对性这个要求”,“情报就是激化了、活化了的的知识”。英国著名学者、情报学家布鲁克斯准确也更科学地描述了情报与知识之间的相互作用关系,他曾提出这样一个基本