

科技信息素养 基础教程

■ 杨云川 杨晶 王清晨 孙蔚 编著

KEJIXINXI SUYANG
JICHU JIAOCHENG



國防工業出版社

National Defense Industry Press

□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □

科技信息素养基础教程

杨云川 杨晶 王清晨 孙蔚 编著



国防工业出版社
National Defense Industry Press

· 北京 ·

内 容 简 介

本书编写贴近网络信息时代,注意吸取信息检索技术最新进展,采用编者部分教学改革成果,将信息素养教育与信息检索相融合,以“实用、好用、够用”为原则,系统介绍了信息素养与信息检索、计算机信息检索概述、中外科技信息资源检索、专利及网络信息资源检索、科技信息综合分析与利用、学术道德和学位论文撰写等内容。

本书可作为高校信息素养教育课通用教材,亦可供其他需要提高自身信息素养的用户学习参考。

图书在版编目(CIP)数据

科技信息素养基础教程/杨云川等编著. --北京:

国防工业出版社,2013. 12

ISBN 978-7-118-09239-4

I. ①科… II. ①杨… III. ①科技情报—情报检索—教材 IV. ①G252.7

中国版本图书馆 CIP 数据核字(2013)第 305154 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

北京奥鑫印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 印张 11 $\frac{3}{4}$ 字数 288 千字

2013 年 12 月第 1 版第 1 次印刷 印数 1—4000 册 定价 28.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777

发行邮购:(010)88540776

发行传真:(010)88540755

发行业务:(010)88540717

前 言

从1946年第一台电子数字计算机的发明算起,当代信息革命已经有67年的历史。这场改变世界的信息革命所带来的信息化,经历了从数字化开始、向网络化和智能化发展的过程。从1984年教育部印发(高教[84]004号)文件算起,“文献检索与利用”作为高校一门公共基础选修课已经走过近30年的历史,课程内容不断更新,信息素养教育伴随科学技术进步,紧跟时代发展步伐。从1994年中国成为国际上第77个正式真正拥有全功能互联网的国家算起,我国用了不到20年的时间,走完了发达国家几十年走过的信息化路程,与很多发达国家一样基本同步进入了一个网络化的时代。

网络化时代的高等学校信息素养教育在各国普遍受到高度重视,2000年1月,美国大学和研究图书馆协会(ACRL)通过了《高等教育信息素养能力标准》;2003年9月,联合国信息素养专家会议发表了《布拉格宣言》,宣称:信息素养是人们有效参与信息社会的一个先决条件,是终身学习的一种基本人权;2008年4月,中国高等学校图书馆情报工作委员会颁发了《高校大学生信息素质指标体系》(讨论稿)。

知识经济和信息网络化时代的一个重要标志是信息急剧增长,知识深度挖掘。“茫茫信息之海洋,浩浩知识之渴望”(We are drowning in information, but starving for knowledge. John Naisbett)。了解信息资源状态,正确识别所需信息,高效获取、分析和充分利用信息已成为当代大学生的一个基本素养。本书旨在适应高等学校“文献检索与利用”课程建设和复合型人才培养的需要,力求帮助学生全面提升科技信息素养和创新意识。

本书注意吸取国内外图书馆学的有关研究成果及科技信息检索技术的最新进展,兼顾省属高等学校数据库资源的实际情况,采用了编者本科信息检索课的部分教学改革成果。编者均为多年从事信息检索课程教学、科研和专业课教学、图书馆情报工作的教师,具有丰富的理论知识基础和实际教学工作经验。本书以“实用、好用、够用”为原则,深入浅出、通俗易懂,避免艰涩,具有如下特色:第一,从内容取舍角度出发,在全面覆盖信息素养能力标准的基础上,突出了易用性和可操作性,增加了中文信息检索平台、专利信息平台的比较,强化了信息道德教育,介绍了学术不端文献检索系统等部分内容;第二,从新技术应用角度出发,不仅更新了文献管理工具和个性化服务等相关内容,而且增加了中文和外文“一站式”检索平台使用方法的介绍;第三,通过应用基于Moodle开发的“信息素养与信息检索”自主学习平台作为辅助教学手段,使采用研究性学习的教学方法进行科技信息素养教育更易实施,通过博客使学生和教师的互动交流更为方便,满足了学生课后乃至毕业后的科技信息检索学习需求。

本书以信息素养教育为主线,以科技信息检索为切入点,力求帮助学生提高信息素养,提高科技信息获取、分析和运用能力,提高知识创新意识。全书共8章,分别为信息素养与信息检索、计算机信息检索、中文网络信息资源检索、外文科技信息资源检索、专利信息资源检索、网络检索工具——搜索引擎、科技信息综合分析与利用以及学术道德和学位论文撰写。其中前言、第6章第3节、第7章、第8章和附录1等由杨云川编写,第3章、第4章和附录2由杨晶编写,第1章和第2章由王清晨编写,第5章和第6章第1节、第2节由孙蔚编写。本书由杨云川、杨晶起草大纲并统稿。

本书可作为高校各专业技术信息检索课的教材用书,也可作为具有科技信息检索需求的科技工作者的参考用书。

本书是在《信息素养与信息检索实用教程》基础上编写的,在编写过程中参考和借鉴了大量文献、资料和网页,引用了部分论点,限于篇幅,仅列出主要参考文献,在此向所有相关作者表示衷心的感谢!

由于学识水平有限以及信息检索资源和技术更新速度很快,书中难免存在疏忽和不足之处,望同行和读者给予指正,以便在下一次修订时加以补充和完善。

编 者

2013年10月22日

目 录

第1章 信息素养与信息检索..... 1

1.1 信息概述..... 1

1.1.1 基本概念 2

1.1.2 信息基本属性 4

1.1.3 信息类型 4

1.2 信息检索..... 7

1.2.1 信息检索内容 8

1.2.2 信息检索基本原理 8

1.2.3 信息检索方法和 手段 9

1.2.4 检索语言与检索效果 评价 9

1.3 信息素养 12

1.3.1 基本概念 13

1.3.2 信息素养教育现状 13

1.3.3 信息素养教育的意义 及发展趋势 16

第2章 计算机信息检索 19

2.1 计算机信息检索概述 19

2.1.1 计算机信息检索 19

2.1.2 计算机信息检索系统 构成 22

2.1.3 计算机信息检索数据库 类型 24

2.2 计算机信息检索基本原理与

技术 25

2.2.1 计算机信息检索 基本原理 25

2.2.2 计算机信息检索 基本技术 25

2.3 计算机信息检索基本步骤与 服务方式 28

2.3.1 计算机信息检索 策略 28

2.3.2 计算机信息检索 基本步骤 28

2.3.3 计算机信息检索 服务方式 30

第3章 中文网络信息资源检索 31

3.1 CNKI 中国知网..... 31

3.1.1 CNKI 中国知网概述..... 31

3.1.2 CNKI 中国知网检索..... 33

3.1.3 CNKI 中国知网个人数字 图书馆 37

3.2 万方数据知识服务平台 40

3.2.1 万方数据知识服务平台 概述 40

3.2.2 万方数据知识服务平台 主要数据库简介 40

3.2.3 万方数据知识服务平台 检索 41

3.3 维普期刊资源整合服务

平台	43	4.1.3 ScienceDirect 数据库 检索	68
3.3.1 维普期刊资源整合服务 平台概述	43	4.1.4 ScienceDirect 数据库 个性化服务	72
3.3.2 维普期刊资源整合服务 平台之期刊文献检索 ..	44	4.1.5 ScienceDirect 免费 资源	73
3.3.3 维普网——仓储式在线 出版平台	46	4.2 EBSCO 网络数据库	73
3.4 中国高等教育文献保障系统 ..	47	4.2.1 EBSCO 网络数据库 概述	73
3.4.1 CALIS 总体资源	47	4.2.2 EBSCO 网络数据库 总体资源	73
3.4.2 CALIS 统一检索 系统	50	4.2.3 EBSCO 网络全文数据库 检索	75
3.4.3 CALIS 联合目录 公共检索系统	50	4.2.4 EBSCO 免费网络 数据库	80
3.4.4 CALIS 其他服务	53	4.3 SpringerLink 平台数据库	81
3.5 国家科技图书文献中心	55	4.3.1 SpringerLink 平台 概述	81
3.5.1 NSTL 总体资源	55	4.3.2 SpringerLink 平台全文 数据库检索	81
3.5.2 NSTL 文献检索	55	4.3.3 SpringerLink 平台个性化 服务	83
3.5.3 NSTL 个性化服务	57	4.4 Wiley Online Library 网络出版 及服务平台	83
3.6 超星网	58	4.4.1 Wiley online Library 网络出版及服务平台 概述	84
3.6.1 超星网概述	58	4.4.2 Wiley Online Library 网络出版及服务平台数据 库检索	84
3.6.2 超星网免费图书 检索	59	4.4.3 Wiley Online Library 网络出版及服务平台 个性化服务	86
3.6.3 超星电子图书 全文检索	59	4.5 ACS 电子期刊平台	86
3.6.4 超星发现	61	4.5.1 ACS 电子期刊平台 概述	86
3.7 读秀学术搜索	64	4.5.2 ACS 电子期刊平台 数据库检索	87
3.7.1 读秀学术搜索概述	64	4.5.3 ACS 电子期刊平台	
3.7.2 读秀学术搜索平台 检索	64		
3.7.3 读秀学术搜索技巧	66		
3.7.4 读秀学术搜索与超星网 比较	66		
第4章 外文科技信息资源检索	67		
4.1 Elsevier - SciVerse 平台	67		
4.1.1 Elsevier - SciVerse 概述	67		
4.1.2 ScienceDirect 数据库 特点	68		

个性化服务	89	6.1.1 搜索引擎发展历史	118
4.6 “一站式”检索	90	6.1.2 搜索引擎类型	119
4.6.1 “一站式”检索概述	90	6.1.3 搜索引擎未来发展趋势	122
4.6.2 Metalib“一站式”检索平台	90	6.2 网络常用学术搜索引擎	123
4.6.3 “一站式”检索发展趋势	91	6.2.1 Google Scholar	123
第5章 专利信息资源检索	93	6.2.2 Scirus 科学信息搜索引擎	124
5.1 知识产权起源和发展	93	6.2.3 OAIster 开放存取搜索引擎/全球联合机构知识库	127
5.1.1 国外知识产权起源和发展概述	93	6.3 网络其他常用搜索引擎	128
5.1.2 中国知识产权起源和发展概述	95	6.3.1 百度	128
5.2 知识产权基本知识	95	6.3.2 Google(谷歌)	130
5.2.1 知识产权概述	96	6.3.3 门户网站搜索引擎	131
5.2.2 专利种类	96	6.3.4 常用中文搜索引擎比较	132
5.2.3 专利性质	97	第7章 科技信息综合分析与利用	134
5.2.4 专利申请	97	7.1 中文信息检索平台比较	134
5.2.5 专利审查与审批	99	7.2 国外检索工具简介	137
5.2.6 专利职务发明与非职务发明	100	7.2.1 工程索引数据库	137
5.3 国际专利分类法	101	7.2.2 科学引文索引数据库	139
5.3.1 IPC 分类体系	101	7.2.3 会议论文引文数据库	143
5.3.2 《IPC 关键词索引》	101	7.3 科技信息搜集与分析整理	144
5.4 中国专利检索网站	102	7.3.1 信息搜集方法与途径	144
5.4.1 中华人民共和国国家知识产权局专利检索网站	102	7.3.2 信息查询步骤	144
5.4.2 中国知识产权网	104	7.3.3 获取原始文献信息途径	145
5.4.3 SIPO 与 CNIPR 比较	109	7.3.4 信息整理方法	145
5.5 国外专利检索网站	110	7.4 科技查新	146
5.5.1 美国专利商标局	110	7.4.1 科技查新发展趋势	146
5.5.2 世界知识产权组织	114	7.4.2 科技查新程序	147
第6章 网络其他学术资源检索工具——搜索引擎	118	7.4.3 科技查新与一般文献检索异同	148
6.1 搜索引擎概述	118		

7.5 个人文献管理软件及其应用	149	必要性	160
7.5.1 NoteExpress 软件	149	8.1.3 学术不端文献检测系统	160
7.5.2 EndNote 软件	151	8.2 学位论文撰写	161
7.5.3 Mendeley 软件	152	8.2.1 学位论文学术水平	161
7.6 基于 Moodle 平台的科技信息素养教育	153	8.2.2 学位论文特点	162
7.6.1 基于 Moodle 构建“信息素养与信息检索自主学习平台”	153	8.2.3 学位论文一般要求	162
7.6.2 “平台”特点	156	8.2.4 学位论文开题	162
7.6.3 信息素养教育模式研究测评分析	156	8.2.5 学位论文中期检查	163
第8章 学术道德和学位论文撰写	158	8.2.6 学位论文撰写	164
8.1 学术道德	158	8.2.7 学位论文答辩	164
8.1.1 学术诚信	158	8.2.8 学位论文撰写基本格式	165
8.1.2 文献信息资源合理使用法律依据和		附录1 《美国高等教育信息素养能力标准》	170
		附录2 计算机实践教学指导	172
		参考文献	179

信息素养与信息检索

随着社会发展和科学技术的进步,人类从原始社会进化到了今天。1980年,未来学家阿尔温·托夫勒(Alvin Toffler)出版《第三次浪潮》,提出“超工业社会”,认为人类社会已经经历了两次浪潮。第一次浪潮是农业革命,由原始狩猎社会变成了农业社会;第二次浪潮是工业革命,把农业社会变成了工业社会,它创造了现有的世界物质文明和精神文明的基础;第三次浪潮是信息革命,它从20世纪50年代中期开始,人类社会从工业社会步入信息社会。在20世纪80年代中期,美国未来学家约翰·奈斯比特(John Naisbitt)提出将人类社会分为农业社会、工业社会、信息社会三个阶段,目前“虽然我们还认为是生活在工业社会,但是事实上已经进入了一个以创造和分配信息为基础的社会”,即信息社会。奈斯比特第一次完整而具体地描述了信息社会的内涵:在新的信息社会中最重要战略资源是信息而不是资本,知识已成为生产力和成就的关键,知识产业为经济社会提供必要的和重要的生产资源。在高技术的信息社会中,人们是利用脑力劳动为主,而不是像工业时代的工人那样从事体力劳动为主。特别是1946年第一台电子数字计算机的发明,使人类更加快速地进入了信息社会。这场改变世界的信息革命所带来的信息化,经历了一个从数字化开始、向着网络化和智能化发展的过程。进入20世纪90年代,随着互联网的迅速发展,以计算机技术、网络技术和通信技术为代表的信息技术的创新性、渗透性、倍增性、带动性促使人类社会形态发生了转变,大部分人从事着信息工作,社会中最重要因素转变为知识,在不知不觉中,人们已经步入了信息社会。

伴随着信息社会的到来,作为这个社会最重要的个体——人类,既离不开社会,又是社会属性的活动体,其本质由各种社会关系决定。如何使每个个体尽早融入、适应快速发展的信息社会,学习能力成为人类必须具备的、最重要的能力之一,具备信息素养以及掌握信息检索技能,逐渐引起社会广泛的关注,成为推进社会进步与发展的重要因素,促进人类发展的必要条件和前提。因而,在大学阶段,应该注意培养自己的信息素养,而信息检索和获取则是信息素养的重要表现之一。

1.1 信息概述

在当今信息社会中,信息无处不在,无时不有,无人不用。信息是一种资源,与知识、文献与情报之间有着密切的联系,是社会进步、经济与科技发展的源泉。信息与空气、水一样重要,它与物质、能源并列构成世界三个要素。每个人都需要从各种各样的信息源中收集和利用信

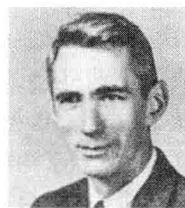
息,为决策提供参考依据。及时获得必要和准确的信息是个人、社会存在与发展的前提条件。通过信息,我们才会考察到人类过去和现在的状态、特征以及发展趋势。因此,我们首先要了解信息、知识、情报、文献的概念及它们之间的关系。

1.1.1 基本概念

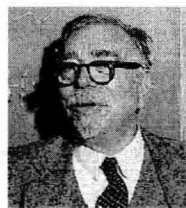
1. 信息含义

信息(Information)有情报、资料、消息、报道之意。在中国历史文献中,“信息”一词最早见于唐代诗人李中的诗词“梦断美人沉信息,目穿长路依楼台”(出自《暮春怀故人》)。《辞海》中对“信息”有两种解释,一是指音讯、消息;二是指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。20世纪中叶以后,信息的概念被引入哲学、信息论、系统论、控制论、传播学、情报学、管理学、通信、计算机科学等领域。所以,站在不同的角度,人们分别从不同的学科给予不同的定义和理解。而今,“信息”不仅早已成为人们使用频率最高的词汇之一,也是人们在日常社会生活、工作、学习和研究中随时随地能感受到和使用的物质属性。

从通信学角度出发,信息是通信的内容。克劳德·艾尔伍德·香农(Claude Elwood Shannon)在《通信的数学理论》中把信息描述为“人们对事物了解不定性的减少和消除,是两次不定性之差。”信息的多少反映了消除的不确定性的多少。



香农



维纳

从控制论角度出发,维纳(N. Wiener)在《信息控制论》一书中指出:“信息是人们在适应外部世界,并使这种适应反作用于外部世界的过程中,同外部世界进行相互交换的内容的总称。”

从认识论角度出发,我国大多数学者认为:信息就是信息,就像“数”的概念一样,它是客观世界中存在的一切物质的一种基本属性,用来提供客观世界的一切存在,从而减少人们对客观世界认识和感知的不确定性。也可以说,它是自然界、人类社会和人类思维活动中普遍存在的一切事物的属性,即自然界和人类社会都会产生信息。人的大脑或感觉器官(或仪器)通过接受外界事物发出的种种消息、指令、数据、信号来识别事物的存在、发展和变化。

从情报学角度出发,《情报与文献工作词汇基本术语》(GB/T 4894—1985)对信息的描述是:“信息是物质存在的一种方式、形态或运动状态,也是事物的一种普遍属性。一般指数据、消息中所包含的意义,可以使消息中所描述的事件的不确定性减少。”这一解释基本涵盖了信息的属性(客观存在性)、作用(消除不确定性)及形式(数据、消息等)。因此,信息是个内涵浅、外延广的概念。

信息是用文字、数据、信号等形式,通过一定的传递和处理,来表现各种相互联系的客观事物在运动变化中所具有特征内容的总称,包括事物内部结构以及外部联系的状态和方式。信息是无形的,存在于整个自然界和人类社会之中。它是客观事物的运动状态和特征的反映,是人们认识事物发展的基础。信息不是物质本身,而是物质的一种基本属性。信息的类型包括自然信息、生物信息、机电信息和社会信息等。

信息素养和信息检索所指的信息,是指经过采集、处理、记录并以可检索的形式存储的数据。

2. 信息与知识

与信息密切相关的另一个概念是知识,即人类认识的成果或结晶,包括经验知识和理论知识。经验知识是知识的初级形态,系统的科学理论是知识的高级形态。世界经济合作与发展

组织(OECD)在1996年的年度报告《以知识为基础的经济》中将知识按内容划分为四类:

- (1) Know-what, 知道什么, 主要是叙述事实方面的知识;
- (2) Know-why, 知道为什么, 主要是指自然原理和规律方面的知识;
- (3) Know-how, 知道怎么做, 主要是指对某些事物的技能和能力;
- (4) Know-who, 知道是谁的知识, 涉及谁知道为什么和谁知道怎么做某些事。

知识是人类对信息、对客观事物规律的认识,它是人们在社会实践中积累起来的经验,是通过实践对客观事物及其运动规律的认识。人们对事物由表及里,由现象到本质,由感性到理性的认识深化,便形成了知识。综上所述,知识是人类主观世界对于客观世界的信息的概括、总结和反映。知识按其获得方式可区分为直接知识和间接知识,按其内容可分为自然科学知识和社会科学知识。知识的总体在社会实践的世代延续中不断积累和发展。

信息是知识的源泉和基础,人们通过信息来认识和掌握自然界、人类社会的活动规律。人脑通过思维重新组合信息,加工信息,使其集合并系统化。人脑对信息加工的成熟度会因人的学术水平和思考时间而存在差异,学术水平高的人对信息精心加工能产生高质量的知识,即信息→知识。人们利用获取的知识创造新信息,即知识→信息。信息与知识密不可分,但信息不等于知识。英国著名情报学家布鲁克斯(B. C. Brooks)这样表述信息与知识的关系:“信息是使人原有的知识结构发生变化的那一小部分知识。”

在大学时代,单纯的学习知识已经远远不能满足时代对个人能力培养的要求,终身学习、终身教育的理念已经被广泛的接受。面对信息的急剧增长、知识的不断更新,我们不得不认真学习和研究获取信息的方法,掌握从大量的、无序的信息和知识中搜索有用的、准确的信息和知识的技能。

3. 知识与文献

为了把人类的知识传播开来和继承下去,人们用文字、图像、符号、音频、视频等手段将其记录下来,写在纸上或存储在介质上,这种附着在各种载体上的记录被统称为文献。我国《国家标准文献著录总则》对文献有如下定义:“文献指记录有知识的一切载体。”人的知识是后天在社会实践中形成的,是对现实的反映。载体是知识的外在表现形式,是文献的外在形态,它可以是甲骨、青铜器、纸介型、胶片、磁带、磁盘、光盘等。知识是文献的内容,知识固化在一定物质载体上,可以在同期社会交流并传递给后人。

文献积累着知识,保存着人类文化遗产,它汇集了人创造的精神财富,是知识的宝库。文献在时间上的传递,实现了人类从古到今的知识的传承和发展;文献在空间上的传播,促进了同时代人们之间知识的交流和沟通,从而推动社会的发展。

4. 信息与情报

“Information”一词即可译为“信息”,也可译为“情报”。在东西方文化的交流过程中,信息和情报两词在媒体上经常混用,但事实上是有区别的。钱学森将情报定义为:“情报是在特定的时间、特定的状态下,传递给特定对象的有用的知识和信息。”也就是说,情报是为解决特定问题所需要的知识或信息。情报来源于知识和信息,具有知识性、信息性、动态性、效用性、时间性等属性。符合人们特定需求的信息和知识一旦进行了有效传递,即具备了价值性、时间性等特征,就成为了情报,而失去时效性的情报又可以还原为知识和信息。知识和信息在人类社会实践活动中被动态地接受与利用,并通过情报的传递功能产生效用。情报在传递过程中启迪人们思维,改变人们知识结构,提高人们的认知能力,并实现知识的社会价值和经济价值。信息与知识、文献、情报之间形成过程的关系如图1-1描述。

宇宙的万事万物都具有自身的信息,所以信息是无限的。人们获取这些信息后,加以分析

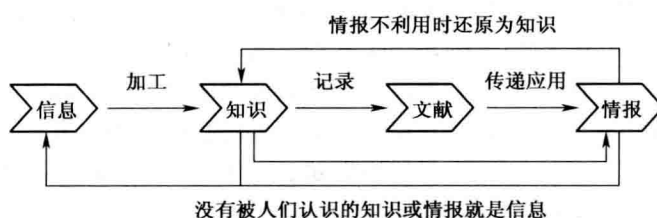


图 1-1 信息、知识、文献与情报关系

研究,从感性认识变成了理性认识,就构成了知识。知识信息用文字、符号、图片、声波、视频等记录在纸张、胶片、磁盘、磁带或者计算机网络上就形成了文献。当人们需要的时候,及时获取了这些知识信息和文献,或者说人们获取这些知识信息和文献在有用的时候,就成为情报。而当情报失去效用时,就又重新回到信息或知识当中。

1.1.2 信息基本属性

信息反映了事物之间的相互关系,是事物的表征,但它并不是事物的本身。信息来源于事物,又必须依赖物质作为传播的载体。信息既与事物有着密切的关系,又有着明显不同的特征。其基本特征有以下几点:

(1) 客观性:信息既不是物质,也不是能量,而是客观事物普遍性的表征,广泛地存在于自然界和人类社会,其存在不以人的意志为转移,是无处不在、无时不有的普遍社会现象。

(2) 存储性:信息存储是针对所采集的信息进行科学有序的存放和保管,以便使用,目的是将杂乱无章的信息集中、整序,提高信息的利用率。

(3) 传递性:信息的传递是与物质和能量的传递同时进行的。信息在事物之间的相互联系必定在信息的流动中发生。信息的传递性表现在人与人之间的消息交换;人与机器、机器与机器之间的信息交换;动物与动物、植物与植物、动物与植物之间的信息交换;人类进化过程的细胞选择也是信息的传递与交换。

(4) 时效性:信息的时效是指从信息发生、接收到利用的时间间隔及效率。信息是有价值的,但是由于事物在不断变化,表征事物存在方式和运动状态的信息也必然会随之改变,即信息本身具有生命周期,其价值会随着时间的推移而改变甚至消失。

(5) 积累性:信息通过人脑思维或人工技术设备的综合、加工和处理,不断积累丰富,使其质量和利用价值不断提高。

(6) 相对价值性:信息的价值不同于普通商品,信息的价值是相对的,完全取决于人们对它的认识和重视程度,也就是说信息的价值更多地体现在人们对它的开发和使用程度上。

(7) 依附性:信息不能独立存在,需要依附于不同的载体。

(8) 可处理性:人脑就是最佳的信息处理器。人脑的思维功能可以对信息进行决策、设计、研究、写作、改进、发明、创造等多种信息处理活动。计算机也具有信息处理功能,可以通过压缩、转换、积累存储在载体上。

(9) 共享性:一般的物质交换遵循等价交换原则,而信息交换双方不会因为交换而失去原有的信息源,相反会为自己增加信息源。信息可以被一个用户使用,也可以被多个用户同时使用,而信息的本质不会被改变。

1.1.3 信息类型

信息与人类智能活动有关的知识、技术、科学、文化、社会等密切联系在一起,其涉及范

围如此广泛,从而形成不同的信息类型,以至于很难用统一的标准进行分类。根据信息内容的特点,大体将其按以下几种情况进行分类:

1. 按信息出版类型划分

1) 图书

图书(Book)是指凡由出版社(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,包括专著、教科书、各种科普读物及专业参考工具书等。图书的内容系统、成熟、定型、可靠性强,是人们从事学习、研究不可缺少的信息来源。不过,传统印刷业图书出版周期较长,体积大,更新速度慢,电子版图书的出现则弥补了这一缺陷。

正规出版的普通图书都有国际标准书号(International Standard Book Number, ISBN 号)。从 2007 年 1 月 1 日起,ISBN 号由 10 位数字升级至 13 位数字组成。

2) 期刊

期刊(Journal)是指一种有固定名称、定期或按宣布的期限出版的连续出版物。与图书相比,期刊最突出的特点是出版数量大、连续性、周期短、内容新,能迅速反映科技研究成果的新信息。期刊作为重要的文献信息源还体现在世界上所有主要检索工具都以期刊为主要收录对象(约占 90% 以上),可以比图书更快、更方便地查到所需资料。每种期刊有一个永久专属的国际标准连续出版物号(International Standard Serial Number, ISSN 号),由 8 位数字组成,分两段,每段四位数。如 ISSN 1002 - 1027, ISSN 1002 - 655X。期刊名称变更,要重新申请 ISSN 号。期刊停刊 ISSN 号不会被其他期刊再使用。

3) 会议文献

会议文献(Proceeding)是指在国内外学术团体举行的专业会议上发表与交流的论文与报告。其特点是专业性强、内容新、学术水平高、出版发行较快,大部分是本学科领域内的新成果、新理论、新方法。与期刊相比,会议文献具有传播情报信息更迅速的功能,能够反映某学科、专业的最新成果和发展水平动向,是科研工作不可缺少的信息源。

4) 学位论文

学位论文(Dissertation)是高等院校或研究机构的学生为取得各级学位,在导师指导下完成的科技研究、科技试验成果的书面报告,须经专家评审、鉴定通过。一般来说具有学术性强的特点,往往有独到的见解。纸质资料一般由学位授予院校保存,电子数据可在网上数据库检索。

5) 专利

专利(Patent)是记录有关发明创造信息的文献,蕴含着技术信息、法律信息和经济信息。广义的专利包括专利申请书、专利说明书、专利公报和专利检索工具,以及与专利有关的一切资料;狭义的专利仅指各国专利局出版的专利说明书。专利文献具有独创性、实用性、新颖性等特征,是重要的技术经济信息来源。

6) 科技报告

科技报告(Report)是指各学术团体、科研机构、大学研究所的研究报告及其研究过程中的记录。科技报告理论性强,是了解某一领域科研进展状况、发展动态的重要情报源。但是科技报告保密性强,难以获取。如美国著名的四大报告:PB 报告、AD 报告、NASA 报告、DOE 报告。

7) 政府出版物

政府出版物(Government)是由政府机构制作出版、或由政府机构编辑并授权指定出版商出版的文献。它主要包括两个大类:一类是行政性文献信息,包括宪法、司法文献;另一类是科

学技术文献信息,主要指政府部门出版的科技报告、标准、专利文献、科技政策文件等。前者约占政府文献信息的60%~70%,后者数量相对较少。

8) 报纸

报纸(Paper)是指出版周期最短的定期连续出版物。报纸的基本特点是内容新、涉及面广、读者多、影响面广。及时性是报纸区别于图书和期刊的最主要特征,又称新闻性和时间性。

9) 技术标准和规范

狭义的标准(Standard)是指按规定程序制订、经公认权威机构批准的一整套在特定范围内必须执行的技术规范、技术标准、操作规程、建议、准则、术语、专有名词等在内的各种技术文件。广义的标准指与标准化工作有关的一切文献,包括标准形成过程中的各种档案、宣传推广标准的手册及其他出版物,报道标准文献信息的目录、索引等。

10) 产品资料

产品资料(Production)是制造厂家和产品销售者介绍其产品的宣传性出版物,包括产品目录、产品说明书、产品样本。通过产品资料可以了解厂家的工艺水平、管理水平和产品发展趋势方面的信息。由于产品资料附有大量图表,因此具有直观的特点。缺点是时间性强,使用寿命短,理论内容介绍极少等。

11) 档案

档案(File)是国家机构、社会组织以及个人从事政治、军事、经济、科学、技术、文化、宗教等活动直接形成的具有保存价值的各种文字、图表、声像等不同形式的历史记录,是完成传达、执行、使用或记录现行使命而留备考查的文件材料。档案以其记录性和原始性于一体的特点而区别于遗留下来的实物,又因其可靠性和稀有性而区别于一次文献,这就使相当一部分档案在一定时间内是受到保护的,在利用上有特殊的要求和价值。其中的技术档案还具有技术性、适用性、保密性等特征。

12) 灰色文献

灰色文献(Gray Literature)是对一些特殊类型的文献信息的总称,一般被看作是非公开出版物。它的研究内容不成熟,流通渠道特殊,没有固定的形态、名称和篇幅,制作份数少,容易绝版,但是往往有特殊的参考价值。

2. 按信息的加工程度(或级别)划分

人们基于信息的传递链和信息的加工处理深度,将信息划分为零次信息、一次信息、二次信息、三次信息等。

(1) 零次信息:是指尚未进入整理或最后定稿的一类信息总称。零次信息是信息的一个部分,是一切信息产生的源信息。

(2) 一次信息:是作者本人以自己的研究或劳动成果为基本素材而创作的原始信息。包括第一次报道的、第一次书写的或第一次出版的那些信息。如期刊论文、会议文献、学位论文等。

(3) 二次信息:是通过对一次信息的加工、提炼和浓缩而形成的信息。它将分散、无序的大量信息转变为有序的、便于管理的信息,成为查找一次信息的工具。如目录、题录、文摘、索引等。

(4) 三次信息:是对一次信息和二次信息进行汇集、综合、分析、评述等深度加工而形成的产物。如字/词典、百科全书、年鉴、手册、综述等。

从零次、一次、二次到三次信息反映了信息的集中和有序化过程。零次信息是一次信息的来源,一次信息是零次信息处理的结果。零次信息可能会被信息获取者直接应用,也可能会经

过一次或多次加工和传递成为一次信息、二次信息、三次信息而被其他信息接受方应用。零次信息、一次信息、二次信息、三次信息基本上后者是在前者的基础上产生的,它们之间存在着“逆向依附性”。

3. 按信息表现形式划分

信息的表现形式繁多,不同形式的信息传递了不同的信息内容。

(1) 文字信息:文字是人们为了实现信息交流、通信联系所创造的一种约定的形象符号。文字、符号、代码均是信息的表述形式,其内容再现于它们的结构属性之中。

(2) 图像信息:图像是一种视觉信息,它比文本信息直接,易于理解。人工创造的图像是抽象或间接的图像信息。随着多媒体技术的发展,各类图像信息库将会极大地丰富人类生活。

(3) 数值数据信息:数值数据是“信息的数字形式”或“数字化的信息形式”。狭义的“数据”是指有一定数值特性的信息,如统计数据、气象数据、测量数据以及计算机中区别于程序的计算数据。广义的数据是指在计算机网络中存储、处理、传输的二进制数字字符编码。文字信息、图像信息、语言信息以及从自然界直接采集的各种自然信息等均可转换为二进制数码,网络中的数据通信、数据处理和数据库等就是广义的数值数据信息。

(4) 话音信息:人讲话实际上是大脑的某种编码形式信息转换成语言信息的输出,是一种最普遍的信息表现形式。音乐也是一种信息形式,是一种特殊的声音信息,它通过演奏的方式表达丰富多彩的信息内容。

4. 按信息载体划分

人类社会发展的不仅仅是信息,记载信息的物质载体形式也在不断更新。总的趋势是由简单到复杂,由低级到高级。信息的载体进化过程分为:

(1) 手写型:包括甲骨、手稿、竹简、帛书/绢书等。

(2) 印刷型:包括纸质图书、期刊、报纸等。

(3) 缩微型:包括缩微胶卷、缩微胶片等。

(4) 机读型:包括电子图书、电子期刊、网络数据库、光盘数据库等。

(5) 声像型:也称视听资料,包括记录有音像的磁性制品。如唱片、录音带、录像带、幻灯片、电影片、多媒体资料等。

1.2 信息检索

信息检索(Information Retrieval,IR)是将信息按一定的方式组织起来,并根据用户的需要找出有关信息的过程和技术。从根本上来说,真正具有信息素养的人不仅要具备信息意识,更要掌握信息检索技术,具备信息获取的能力。根据美国科学基金委员会、凯斯工学院研究基金会以及日本国家统计局的初步统计,科研人员在—个研究项目中,其时间分配大致如图 1-2 所示。

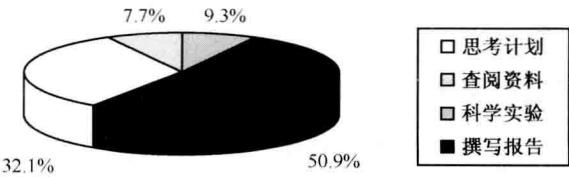


图 1-2 科研人员研究时间分配图

从图 1-2 可见,在整个科研的时间分配比例上,用于查找和阅读情报资料的时间达 50.9%,可见信息之重要,信息检索能力之重要。需要注意的是信息检索(Information Retrieval)与信息搜索(Information Search)是两个不同的概念,两者之间的差异如表 1-1 所列。

本节主要介绍的是信息检索。

表 1-1 信息检索和信息搜索的区别

	信息检索	信息搜索
获取途径	是从有序信息集合中识别与获取所需信息	从任意资源获取所需信息
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、学习
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

1.2.1 信息检索内容

信息检索可以从广义和狭义两个角度理解。从广义上讲,信息检索包括两个过程:①信息存储(Information Storage),指将杂乱无序的信息有序化存储形成信息集合的过程(即信息的标引、加工和存储过程);②信息检索,即用户根据需要从信息集合中找出特定信息的过程。

从狭义上讲,信息检索通常指第二个过程,仅指信息检索这一过程。

信息检索内容主要包含六种形式:

(1) 文献检索(Document Retrieval):是以文献(包括目录、题录、文摘、索引和全文)为检索对象,以科学的方法利用专门的工具,从大量的文献资料中迅速、准确、完整地查找到文献资料的过程。

(2) 数据检索(Data Retrieval):是以数值或数据(包括数据、图表、公式等)为对象的检索。它一般以数据大全、手册、年鉴等为检索工具。

(3) 事实检索(Fact Retrieval):是以某一客观事实为检索对象,查找某一事物发生的时间、地点及过程的检索。一般利用字词典、年鉴、百科全书、手册等作为检索工具。

(4) 全文检索(Full Retrieval):是计算机程序通过扫描文章中的每一个词,对其建立一个索引,并指明该词在文章中出现的次数和位置,当用户查询时根据所建立的索引进行检索。

(5) 超文本检索(Hypertext Retrieval):是利用非线性信息组织方法建立数据库并检索相关信息的方式。超文本的基本组成元素是节点和节点间的逻辑连接链。节点存储信息,节点中的信息和逻辑连接链被联系在一起,构成相互交叉的信息网络,当用户激活某个节点时,计算机就会在屏幕上显示出相应节点的信息。

(6) 超媒体检索(Media Retrieval):是在超文本的基础上发展而来的。随着多媒体技术的发展,开始将图像、图形视频、动画、声频等超媒体信息载体用来建立超链接。存储对象从单维发展到多维,存储空间不断扩大。

1.2.2 信息检索基本原理

信息检索的基本原理就是将用户需求(检索)标识和信息集合中存储(标引)标识进行比较和选择,其核心是两者通过检索语言进行匹配(Match)的过程。两者标识匹配,则具有这些特征标识的信息就从检索系统输出,输出的信息就是检索要求的信息。匹配有多种形式,既可