



高等职业教育“十二五”规划教材

高职信息检索 实用教程

Gaozhi Xinxì Jiansuo Shiyong Jiaocheng

◎ 主编 周 云



配电子课件



机械工业出版社
CHINA MACHINE PRESS

高等职业教育“十二五”规划教材

高职信息检索实用教程

主 编 周 云

副主编 谷松立 邵莉娟

参 编 邵黄芳 丁海晖



机械工业出版社

本书结合工程类高职高专院校工学结合的特点,对信息检索工具的功能、类型、特点进行介绍,并侧重对信息检索途径、检索工具使用方法和基本步骤作了系统阐述,试图体现本书的实用性;同时为了让信息检索技能更好地与高职高专学生的毕业综合实践成果相匹配,就毕业顶岗实习报告和毕业设计(论文)的相关模式、资料收集和整理要点及撰写步骤、方法以及对信息的综合利用等进行了介绍。因此,本书不仅可作为工程类高职高专学生信息检索课程的教材和课外参考书,也可为工程专业技术人员获取相关信息提供帮助。

为方便教学,本书配备电子课件等教学资源。凡选用本书作为教材的教师均可登录机械工业出版社教材服务网 www.cmpedu.com 免费下载。如有问题请致信 cmpgaozhi@sina.com,或致电 010-88379375 联系营销人员。

图书在版编目(CIP)数据

高职信息检索实用教程 / 周云主编. —北京:机械工业出版社, 2013.6

高等职业教育“十二五”规划教材

ISBN 978-7-111-42479-6

I. ①高… II. ①周… III. ①情报检索—高等职业教育—教材 IV. ①G354

中国版本图书馆 CIP 数据核字(2013)第 098353 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

策划编辑:王玉鑫 责任编辑:王玉鑫 孙晶晶

责任校对:肖琳 封面设计:张静

责任印制:杨曦

北京双青印刷厂印刷

2013 年 6 月第 1 版第 1 次印刷

184mm×260mm·15.75 印张·387 千字

0 001—3 000 册

标准书号:ISBN 978-7-111-42479-6

定价:30.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

电话服务 网络服务

社服务中心:(010) 88361066

教材网:<http://www.cmpedu.com>

销售一部:(010) 68326294

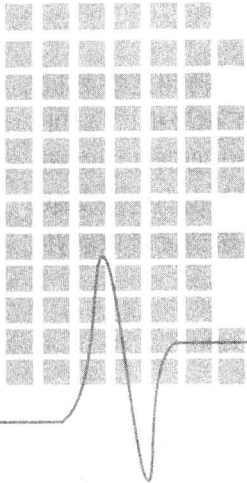
机工官网:<http://www.cmpbook.com>

销售二部:(010) 88379649

机工官博:<http://weibo.com/cmp1952>

读者购书热线:(010) 88379203

封面无防伪标均为盗版



前言

PREFACE

本书主要是为了培养高职高专院校学生和工程技术人员的信息素养能力而编写的。在瞬息万变的数字化时代,信息获取能力已成为新的核心竞争力。高职高专院校培养的学生作为未来信息社会中信息利用的主要群体之一,欲使他们成为高素质人才并赋予他们终身学习的能力,高职高专教育必须积极面对社会变革的严峻挑战,将提高学生信息素养作为人文素质教育的重要目标和内容,将信息素养教育作为高职高专教育教学改革的重要任务之一。

自我国教育部 1984 年颁发了《关于在高等学校开设<文献检索与利用课>的意见》的文件后,各院校积极制定并实施自己的信息素养教育计划和课程,已相继出版了多种与高职高专教育匹配的信息检索和利用方面的教材。随着科学知识更新周期的不断缩短以及信息技术的飞速发展,尤其是目前我国高职高专教育强调把终身教育放在社会教育的中心地位,形成学习型社会以此完善现代职业教育体系。为此本书强化数字化信息素养教育与专业教育的有机结合,是一本融入最新信息技术、嵌入专业教学、体现高职高专教育特色的教材。在编写形式上以提高学生实践检索技能为主要目的,注重“自主+协作”学习环境的构建,每章配有习题,可采取教师讲授为辅,学生练习为主的教学模式。并配有问卷调查表,注重课程教学效果的反馈。

全书共 8 章,分四部分,第一部分(第 1、2 章)主要介绍了信息检索的基本知识和基本原理,由周云编写;第二部分(第 3 章)主要介绍了图书馆的分类和功能,包括图书馆的产生、发展、类型和职能以及对图书馆的图书分类和目录组织,由邵莉娟编写;第三部分(第 4~6 章)主要介绍以计算机为主的各种信息检索,包括纸质工具书和网络工具书的主要检索方法和技巧,在介绍计算机信息检索的原理、检索的基本程序以及检索的基本策略的基础上,重点介绍使用最广泛的中外数据库检索方法,并对特种文献检索(专利、标准文献、科技报告、学位论文等)专门作了说明,由谷松立、邵莉娟编写;第四部分(第 7、8 章)主要介绍如何分析收集的信息并进行预测,如何选取有价值的课题以及相关文献综述的写作技巧,同时结合高职高专学生毕业综合实践成果,介绍了毕业顶岗实习报告和毕业设计(论文)的相关模本、资料收集和整理要点及撰写步骤、方法等以及提升对信息的综合利用能力,诸如如何获取就业信息,如何策划校园活动,如何考取与自己专业有关的相关证书等,由周云编写。邵黄芳、丁海晖提供有关案例和部分习题。



本书编写承蒙浙江大学研究馆员章云兰老师、杭州电子科技大学研究馆员陈琳老师和机械工业出版社的大力支持和协助，在编写过程中又遇上浙江省高校图工委正积极推广 ZADL（浙江省高校数字图书馆）的时机，因此得到浙江省高校数字图书馆滨江分中心浙江中医药大学图书馆的帮助，在此表示衷心感谢。同时编写中又参阅了大量文献，在此也向这些文献的作者致以诚挚谢意。由于编者水平有限，书中难免存在遗漏和不妥之处，敬请读者谅解并加以批评指正。

编者

2012年11月25日

本书编写承蒙浙江大学研究馆员章云兰老师、杭州电子科技大学研究馆员陈琳老师和机械工业出版社的大力支持和协助，在编写过程中又遇上浙江省高校图工委正积极推广 ZADL（浙江省高校数字图书馆）的时机，因此得到浙江省高校数字图书馆滨江分中心浙江中医药大学图书馆的帮助，在此表示衷心感谢。同时编写中又参阅了大量文献，在此也向这些文献的作者致以诚挚谢意。由于编者水平有限，书中难免存在遗漏和不妥之处，敬请读者谅解并加以批评指正。

本书编写承蒙浙江大学研究馆员章云兰老师、杭州电子科技大学研究馆员陈琳老师和机械工业出版社的大力支持和协助，在编写过程中又遇上浙江省高校图工委正积极推广 ZADL（浙江省高校数字图书馆）的时机，因此得到浙江省高校数字图书馆滨江分中心浙江中医药大学图书馆的帮助，在此表示衷心感谢。同时编写中又参阅了大量文献，在此也向这些文献的作者致以诚挚谢意。由于编者水平有限，书中难免存在遗漏和不妥之处，敬请读者谅解并加以批评指正。

目 录

CONTENT

前言

第1章 信息检索概论1

- 1.1 信息概述1
- 1.2 信息检索6
- 1.3 信息素养12
- 本章小结19
- 练习题19

第2章 信息检索基本理论20

- 2.1 信息检索工具20
- 2.2 信息检索语言22
- 2.3 信息检索的途径、方法与步骤27
- 本章小结31
- 练习题31

第3章 图书馆的使用32

- 3.1 图书馆概述32
- 3.2 图书馆图书分类与目录组织39
- 本章小结47
- 练习题48

第4章 参考工具书检索49

- 4.1 工具书概述49
- 4.2 常用工具书手工检索54
- 4.3 网络工具书67
- 本章小结75
- 练习题75

第5章 计算机网络信息检索76

- 5.1 计算机信息检索概述76
- 5.2 图书馆中文数字资源检索86
- 5.3 网络资源查询101
- 本章小结118
- 练习题118

第6章 特种文献检索120

- 6.1 专利文献及其检索120
- 6.2 标准文献及其检索139
- 6.3 科技报告及其检索147
- 6.4 学位论文及其检索152
- 6.5 会议文献及其检索154
- 6.6 政府出版物156
- 6.7 科技档案158
- 6.8 产品资料160
- 本章小结163
- 练习题163

第7章 信息分析与毕业实习报告 撰写165

- 7.1 信息分析165
- 7.2 文献综述171
- 7.3 高职生毕业综合实践成果指导178
- 本章小结208
- 练习题208

第8章 信息资源综合利用209

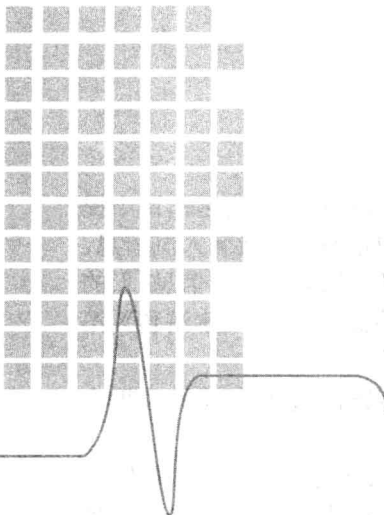
- 8.1 就业材料准备209
- 8.2 校园活动方案制订212
- 8.3 职业资格证书215
- 本章小结233
- 练习题233

综合练习234

“信息检索与利用”课程问卷调查234

附录 高职高专毕业设计(论文)抽评 参考标准238

参考文献240



第 1 章

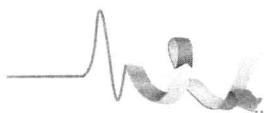
信息检索概论

学习目的：信息素养是一种特殊的、涉及知识面很宽的综合能力，它包含人文的、技术的、经济的、法律的诸多因素，和许多学科有着紧密的联系。信息技术支持信息素养，通晓信息技术强调对技术的理解、认识和使用技能。信息素养的重点是信息内容的传播、分析，包括信息检索以及评价等方面，它是一种既需要通过熟练的信息技术，也需要通过完善的调查方法，经过鉴别和推理来完成了解、搜集、评估和利用信息的知识结构。那么，什么是信息，什么是信息素养，什么是信息检索，大学生如何提高信息素养。

1.1 信息概述

20 世纪 80 年代以来，由于第三次信息技术革命的推动和知识经济的勃兴，人类逐步发展到信息社会阶段。回顾 30 多年的发展历程，信息革命带来的是全方位的变革。当今时代一切事物都无时无刻不处于快速变化之中。以知识经济为显著特征的信息社会，已经成为社会发展的一种无法抗拒的趋势。据专家估计，20 世纪 40 年代以来，产生和累计的信息量已大大超过了在此之前人类有史以来的所有信息量之和。19 世纪以来人类知识信息量每 50 年增长一倍；20 世纪中叶每 10 年增长一倍，70 年代以后每 5 年增长一倍。在面对呈几何级数增长的信息数量时，许多国家在以“国情咨文”、“白皮书”、“蓝皮书”等形式表述对 21 世纪高等教育发展趋势关注中，特别强调，在人才培养上，要更加注重素质和能力的培养，特别是创新能力、信息能力的培养。

信息素养已成为人们特别是大学生必须具备的一种基本能力。信息素养(Information Literacy)更确切名称应该是信息文化(Information Literacy)。信息素养是一种对信息社会的适应能力。美国教育技术 CEO 论坛 2001 年第 4 季度报告提出 21 世纪的能力素质，包括基本学习技能(指读、写、算)、信息素养、创新思维能力、人际交往与合作精神、实践能力。信息素养是其中一个方面，它涉及信息的意识、信息的能力和信息的应用。



1.1.1 什么是信息

1. 关于“信息”一词

在汉语中,“信息”一词最早出现在 1000 多年前的唐诗中:诗人杜牧在《寄远》诗中提到“塞外音书无信息,道傍车马起尘埃。”;李中的《暮春怀故人》中也有“梦断美人沉信息,目穿长路倚楼台。”的佳句。宋代的李清照则发出“不乞隋珠与和璧,只乞乡关新信息。”的感叹。在她心目中,来自家乡的信息比珍贵的“隋珠”与“和璧”的价值更高。《水浒传》第四十四回:“宋江大喜,说道:‘只有贤弟去得快,旬日便知信息。’”

在古希腊文中,“信息”单词为 $\mu\omicron\rho\phi\eta$ 或 $\epsilon\acute{\iota}\delta\omicron\varsigma$ (注: $\epsilon\acute{\iota}\delta\omicron\varsigma$ 为著名哲学家柏拉图和亚里士多德经常用的词),以表示理想的身份或事物的本质。

在英语中,表达“信息”这个概念的词是 information,来源于拉丁文宾格形式 (informationem) 的主格 (informatio),这个名词是由动词反过来又衍生出 “informare” (告知),“to give form to the mind”, “to discipline”, “instruct”, “teach”。

人对世界的认知和改造过程就是收集信息、加工信息和传递信息的过程。人们通过电视、电话、报纸、杂志等各种媒体,无时无刻不在收集、加工和传递、利用着大量的信息,如通过天气预报获取气象信息,可以合理地安排生产、生活。在行政工作中,看材料、学文件是获取信息,作决策、批文件是处理信息,作指示是传递信息。可见,信息来源于客观世界,范围广大,具有一定的利用价值,可以通过载体为人们所获知,用来指导人类认识世界、改造世界。

2. 信息的含义

近几十年来,信息的含义问题,是信息科学等多学科、多领域中研究的一个基本问题。但至今信息还没有一个公认的定义。

美国著名学者 F. 马克卢普 (Machlup) 等人在 20 世纪 80 年代前期曾经仔细评价了关于信息的各种不同含义,并指出对信息的概念存在着多种不同的理解。在“信息”一词的广泛应用中,比较普遍的看法有以下 6 种:

- 1) 信息是事实。
- 2) 信息是数据。
- 3) 信息是沟通 (传播)。
- 4) 信息作为一种能量形式。
- 5) 信息作为一种商品。
- 6) 信息作为知识。

20 世纪 80 年代末,美国著名的信息系统专家 A. 德本斯 (Debons) 等人在其获奖学术著作《信息学:一个综合的视野》中,提出从人的整个认知过程的动态连续体中理解信息的重要观点。他们将认知过程表达为:事件—符号—数据—信息—知识—智慧。这个连续统一体中的任一组成部分,都产生于它的前一过程,例如,“信息”是源于“数据”的,又是“知识”的来源。

对所说的 4 个主要概念:数据、信息、知识和智慧,其含义进一步概括如下:

数据 (Data) —— 字母、数字、线条和符号等,用于表达事件和它们的状态,并根据正式的规则和惯例加以组织。

信息 (Information) —— 对感觉 (被告知) 的某种认知,并给出物理形式 (数据) 的表



述,其表述形式有利于知识的处理。

知识(Knowledge)——超越感知的认知状态。知识意味着主动参与、理解和扩展人的理解水平以处理非常规问题。知识也是指对人的经验的有组织的记录,并赋予物的表达形式(如书籍、报告)。

智慧(Wisdom)——人进行判断时对知识的运用,通常与文化或社会普遍认同的准则或价值观念有关。

可以看出,在德本斯关于信息含义的解释中,有两点是被强调的,即:一是被告知的状态;二是对刺激感官的数据产生某种意识。

在多种不同学科的研究中,信息论的奠基人 C. E. 香农(Shannon)在 1948 年对信息的含义所提出的观点仍然是最基本且具有普遍意义的,香农认为,“信息就是不定性的排除”。这一科学、抽象的论述有助于人们对各种不同领域的信息活动过程的理解。

1.1.2 与信息相关的几个概念

1. 文献

在我国,“文献”一词最早出现于春秋战国时期的《论语·八佾》中。孔子说:“夏礼吾能言之,杞不足征也;殷礼吾能言之,宋不足征也;文献不足故也,足,则吾能征之矣。”这段话的意思是:夏代的制度我知道,杞国的我不知道;殷代的制度我知道,宋国的我就不了解;这是文献不足的缘故,如果足,那么我就能知道了。汉末建安时期玄学大师何晏在其《论语集解》中引汉代经学大师郑玄注:“献,犹贤也”,至宋代,朱熹在《四书章句集注》中沿袭郑玄、何晏的训诂,进一步明确注释为:“文,典籍也;献,贤也。”在此,“文”是指典章制度的文字资料,“献”是指见多识广、熟悉掌故的人。我国古代早期的文献概念基本上是广义的。

宋末元初,著名史学家、文献学家马端临编著了我国历史上第三部记述“典章经制”的通代政书,并以“文献”命名其书为《文献通考》。马端临在其书“自序”中对“文献”解释为“信而有征”,这当属“文献”概念的狭义性定义。明代编成的我国历史上规模最大的类书《永乐大典》,其书初名曾为《文献集成》,以《永乐大典》所收类书的综合性与广泛性而论,此所谓“文献”当属广义。

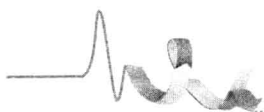
随着时间的推移,文献的概念伴随着人类对于社会的认识而不断演化,现代人们对文献概念的诠释是用文字、图形、符号、声频、视频等技术手段记录人类知识的一种载体,或理解为固化在一定物质载体上的知识。也可以理解为古今一切社会史料的总称。现在通常将文献理解为图书、期刊等各种出版物的总和。文献是记录、积累、传播和继承知识的最有效手段,是人类社会活动中获取情报的最基本、最主要的来源,也是交流信息和传播情报的最基本手段。正因为如此,人们把文献称为情报工作的物质基础。

2. 知识

古往今来,人们对知识的理解是仁者见仁,智者见智。最早对它下定义的是古希腊的哲学家柏拉图,他把知识定义为“证明(Justified)了的真信念”,这个定义直到 20 世纪 60 年代仍然被广泛接受,成为知识论的一个基本概念。

当代著名的认知心理学家皮亚杰认为:“知识是主体与环境或思维与客体相互交换而导致的知觉建构,知识不是客体的副本,也不是由主体决定的先验意识。”

20 世纪 60 年代开始,西方对于“知识”的概念有一个较流行的说法——知识就是 3W1H,



即知道是什么 (Know What)、知道为什么 (Know Why)、知道怎么做 (Know How)、知道是谁 (Know Who)。在我国,这一定义也被人们广泛接受。这表明“知识”这一概念的外延是极其宽广的。

《现代汉语词典》(第6版)对“知识”的释义之一是:“人们在社会实践中所获得的认识和经验的总和。”从本质上说,知识属于认知的范畴。社会实践是在有知识的条件下进行的,是知识所产生的活动,如果没有知识作前提,则无法进行社会实践。

美国社会学家丹尼尔·贝尔对知识的定义是:对事实和思想的一套系统的阐述所提出的合理性判断或经验的结果。而马克卢普 (Machlup) 在《美国的知识生产与分配》一书中认为,知识就是根据已认识的事物所作的客观解释。

知识的定义纷繁复杂,日常语言中使用的知识概念和哲学探讨知识本质问题时的“知识”概念是有区别的,前者的外延要比后者广泛得多。对知识的认识,一直是人们感兴趣的问题。因此,关于知识的定义有上百个,但没有一个能从本质上区分知识与非知识的界限。

知识是人类社会实践经验的总结,是人的主观世界对于客观世界的概括和真实反映。知识是人类通过信息对自然界、人类社会以及思维方式与运动规律的认识,是人的大脑通过思维重新组合的系统化的信息集合。因此,人类不仅要通过信息感知世界、认识和改造世界,而且要根据所获得的信息组成知识。可见,知识是信息的一部分,文献是信息的一种载体。文献是信息传递的主要物质形式,是吸收、利用信息的主要手段。

知识 (Knowledge) 是人类通过信息对自然界、人类社会及思维方式与运动规律的认识与概括,是人的大脑通过思维重新组合的系统化的信息,是信息中最有价值的部分。信息是创造知识的原材料,知识是信息加工的抽象产物。根据知识能否清晰地表述和有效转移,可以把知识分为主观知识(又称为隐性知识, Tacit Knowledge)和客观知识(又称为显性知识, Explicit Knowledge)。显性知识,也称编码知识,是指能用文字和数字表达出来,人们可以通过口头传授、教科书、参考资料、期刊杂志、专利文献、视听媒体、软件和数据库等方式获取,“容易以硬数据的形式交流和共享,并且经编辑整理的程序或者普遍原则”。隐性知识是指高度个性化而且难于格式化的知识,它存在于人脑之中,是个人的见解,只有通过个人经验才能获取,主观的理解、直觉和预感都属于这一类。当隐性知识通过某种载体记录下来,就可打破时空,成为可传递的显性知识。

3. 情报

情报 (Intelligence) 是指被传递的知识或事实,是知识的激活,是运用一定的媒体(载体),越过空间和时间传递给特定用户,解决科研、生产中的具体问题所需要的特定知识和信息。情报自产生开始就带有鲜明的定向性。“情报”一词在我国最初的含义多与战事相关。其典型的定义是:“战时关于敌情之报告,曰情报”。在日本人森欧的译著《战争论》中,情报是指“有关敌方或敌国的全部知识”。当然,这一带有军事特征的情报定义,具有明显的战争年代的特征。

在现代,学术界对于情报的理解存在认识上的共性。其一,情报来自知识,来自于对知识的加工处理;其二,情报不等同于广义的知识,而只是“作为交流对象的有用知识”。现代的情报概念,已经延伸至“特定性”情报、“决策性”情报和“竞争性”情报等,进入了社会各阶层、各领域。

情报具有目的性、特定性的观点已为人们所普遍接受,也就是说,信息、知识成为情报的首要条件是要有社会的需求,这种需求被社会情报系统接收后,就由专门的情报人员对其



加以分析、研究,运用已有知识对相关信息进行加工、整理,其结果产生情报,再由社会情报系统传递出去,为社会所利用,并产生一定的服务效果。社会在利用情报的过程中,可能产生新的需求,同时,也可能产生新信息。

4. 信息、知识与情报的逻辑关系

知识、情报都是特殊的信息。从逻辑上看,信息、知识与情报三者的概念之间是相容关系。具体来说,信息与知识、情报之间是属种关系,信息是属概念,知识、情报是信息之下具有交叉关系的种概念。用它们的外延集合来表达这种关系,可以得到信息、知识与情报的逻辑关系图,如图 1-1 所示。

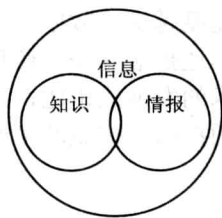


图 1-1 信息、知识与情报的逻辑关系

1.1.3 信息的特征

尽管从不同的角度出发对信息有不同的定义,但是信息的一些基本性质还得到了共识。

1. 普遍性

信息的普遍性是指信息是无处不在的。信息普遍存在于自然界、人类社会之中,也存在于人类的思维或精神领域之中。同客观世界所产生的信息一样,精神世界、思维领域的信息也具有相对独立性,可以被记录下来加以保存、复制或重现。

2. 客观性

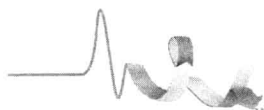
信息不是物质,而是物质的产物,即先有信息反映的对象,然后才有信息。无论借助于何种载体,信息都不会改变其所反映对象的属性,因此,信息具有客观性。例如,气象预报无论是通过广播、电视、报纸,还是通过其他载体,反映的都是自然界的客观变化。

3. 可扩充性与可压缩性

一切领域都有信息,随着时间的推移和事物的运动、发展、变化,信息经过不断地开发利用,会扩充、增值,成为取之不尽、用之不竭的资源。同时,经过加工整理,又可使之精练、浓缩,将信息内容物化在不同的物质载体上,因此,信息又有可压缩性。信息的可压缩性还表现在信息储存上,储存的形式多种多样,如人和动物的记忆就是其中的一种。

4. 时效性

信息的时效性是信息的重要特征,是指信息从发出、接收到进入利用的时间间隔及其效率。信息的时效性与信息的价值性密不可分。任何有价值的信息,都是在一定条件下起作用的,如时间、地点、事件等,离开一定的条件,信息将会失去应有的价值。从某种意义上讲,信息的价值取决于信息的时效性,特别是反映客观事物某种发展趋势的信息,时效性越强,信息的价值越大;反之,信息就会失去作用。因此,信息价值的大小取决于信息的时效性。



5. 可替代性

信息的可替代性有两方面含义：一方面是指信息的物质载体形态是可以相互替代的，如语言信息经过记录变成文字信息，即文字信息替代了语言信息；另一方面是指信息的利用可以替代资本、劳动力和物质资料，这一点在经济学上的作用尤其显著。管理学认为，信息是管理的重要手段和工具，正确利用信息是提高管理水平的重要环节，利用好信息，就可以代替资本和物质的投入。

6. 可传递性与可扩散性

信息的可传递性是指信息可以借助一定的物质载体传递给感受者、接收者的特性。信息可以进行空间和时间上的传输，传输速度越快，效用就越大。科技的发展，使传播信息的网络覆盖面越来越大，从而使信息得以迅速扩散开来。信息的可扩散性与信息传递技术的发展密切相关，信息的扩散速度与传递技术的发展成正比，即传递技术发展得越快，信息扩散的速度越快。随着信息传播手段和技术的提高，信息的扩散性已表现得越来越突出。

7. 共享性

共享性是指信息能够同时为多个使用者所利用，信息扩散后，信息载体本身所含的信息量并没有减少，这是信息与实物、能量等的根本区别。通过传递，信息迅速为大多数人所接收、掌握和利用，并会产生出巨大的社会效应。正因为信息的这一特性，社会才为保护信息开发者的合法权益，补偿其在开发、整理某些信息过程中付出的代价，制定了知识产权制度。

1.2 信息检索

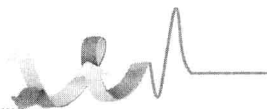
1.2.1 信息检索的概念及类型

1. 信息检索的概念

信息检索 (Information Retrieval)，又叫做情报检索或文献检索，是指将信息按一定的方式组织、存储起来，并根据信息用户的信息需求查找所需信息的过程。从广义上讲，信息检索包含信息的存储和检索两部分。信息的存储主要是指对在一定范围内的信息选择的基础上进行信息特征描述、加工并使其有序化，即建立数据库。信息的检索是借助一定的设备和工具，采取一系列方法与策略，从数据库中查找所需信息。通常所说的信息检索是指狭义概念的信息检索，即从信息集合（数据库）中查找所需信息的过程，也就是信息查寻 (Information Search 或 Information Seek)。

2. 信息检索的起源

信息检索起源于图书馆的参考咨询和文摘索引工作，从 19 世纪下半叶开始发展，至 20 世纪 40 年代，索引和检索已成为图书馆独立的工具和用户服务项目。随着 1946 年世界上第一台电子计算机问世，计算机技术逐步进入信息检索领域，并与信息检索理论紧密结合起来；脱机批量情报检索系统、联机实时情报检索系统相继研制成功并商业化，20 世纪 60 年代到 80 年代，在信息处理技术、通信技术、计算机和数据库技术的推动下，信息检索在教育、军事和商业等各领域高速发展，得到了广泛的应用。Dialog 国际联机情报检索系统是这一时期



的信息检索领域的代表,至今仍是世界上最著名的系统之一。随着计算机静、动态图像信息和声音信息处理技术等先进科研成果的应用,现代信息检索范围得到了拓展,检索功能得到了增强,检索效率得到了提高。

3. 信息检索的类型

(1) 根据信息存储和检索内容的不同,信息检索可分为文献检索、数据检索和事实检索三种方式。

1) 文献检索。文献检索是指查找用户所需文献的线索或者原文的检索。如查找某一主题的相关文献,对某研究课题立项的文献查新,或从事新产品开发时需要查找有关最新研究动态等。文献检索是一种相关性检索,检索结果是某一专题的文献线索(文摘、题录),一般要经过阅读文摘后才能决定取舍。文献检索主要是利用二次文献进行,如各种载体形式的目录、题录、文摘、索引等。文献检索是信息检索中最基本、最重要的类型。

2) 数据检索。数据检索是指查找用户所需特定数据的检索。例如,查找各种统计数据、某一公式、图表、化学分子式、设备型号、技术参数等。数据检索是一种确定性检索,主要是利用各种词典、手册、百科全书、年鉴等参考工具书进行,也可以利用各种参考型数据库进行检索。

3) 事实检索。事实检索是指以特定的事实为检索对象的一种检索。例如,查找某一人、机构的基本情况;某一事件发生的时间、地点、过程等。事实检索和数据检索一样,也是一种确定性检索,不同的是需要对检索出来的数据进行逻辑推理后,方可得出结论。可利用百科全书、手册、年鉴、名录及相关数据库等参考工具进行检索。

(2) 按检索方式区分,信息检索可分为手工信息检索和机器信息检索。

1) 手工信息检索。手工信息检索以手工操作的方式,利用检索工具书进行信息检索。其优点为直观、灵活,便于控制检索的准确性。

2) 机器信息检索。机器信息检索是指计算机信息检索,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。它目前已成为主流方式。

(3) 按检索要求区分,信息检索可分为强相关检索和弱相关检索。

1) 强相关检索。强相关检索强调检索的准确性,向用户提供高度对口信息的检索,也称为特性检索。

2) 弱相关检索。弱相关检索强调检索的全面性,向用户提供系统、完整的信息检索,也称为族性检索。

(4) 按检索的事件跨度区分,信息检索可分为定题检索和回溯检索。

1) 定题检索。定题检索是指查找有关特定主题最新信息的检索,又称 SDI 检索。特点是只检索最新的信息,时间跨度小。该检索在文献信息库更新时运行,适合信息跟踪,便于及时了解有关主题领域的最新发展动态。

2) 回溯检索。回溯检索是指查找一段时间内有关特定主题信息的检索,也称为追溯检索。其特点是既可查找过去某一段时间的特定主题信息,也可以查找最近的特定主题信息。这是用户利用最多的检索方式。

(5) 按检索对象的形式区分,信息检索分为文本检索和多媒体检索;按检索对象的信息组织方式区分,信息检索分为全文检索、超文本检索和超媒体检索;按检索途径的特点区分,信息检索分为常用法、回溯法和循环法。



1.2.2 信息检索的意义

1. 避免重复劳动

科学研究具有继承和创造两重性,科研的两重性要求科研人员尽可能多地占有相关资料、情报。从实践经验看,科研中出现的绝大多数问题都有必要而且有可能通过查找科技文献得到启发甚至得到解决。可以说,一项科研成果中 95% 是别人的, 5% 是个人创造的。因此研究人员在开始着手研究一项课题前,必须利用科学的信息检索方法来了解这个课题的情况,即前人在这方面做过哪些工作,还存在什么问题,以及相邻学科的发展对研究这项课题提供了哪些新的有利条件等与研究课题有关的科技信息,只有这样,才能正确地制订研究方案,防止重复研究,并少走弯路。在我国,重复研究的现象比较严重。一方面,重复研究国外已有的技术,另一方面,国内各机构之间相互重复研究,因此,只有研究人员加强科技文献检索意识和能力的培养,才能彻底改变这种状况。

2. 提高科研效率

现代文献的数量和类型增长十分迅速。实际上,任何人都不可能将世界上所有的文献都阅读完。据调查统计,科研人员查阅文献的时间占 40%~50%,如果利用科学的信息检索方法,就可以缩短研究人员查阅文献的时间,提高科研效率。

3. 促进专业学习

科技文献检索的学习将把学生引导到超越教学大纲的更广的知识范围中去。一个学生在大学学习中已获得进行科研的基本知识,但在校学习时间毕竟有限,参加工作以后,仍需不断更新知识,才能适应科技的迅速发展。掌握了信息检索的方法与技能,就能很快找到一条吸取和利用大量新鲜知识的捷径,进入旺盛的创造期。

1.2.3 信息检索对象——信息资源

1. 信息资源的概念

可以从广义和狭义两种角度来认识和理解信息资源的含义:广义信息资源是指人类社会经济活动中经过加工处理的、有序化并大量累积后的有用信息的集合。狭义信息资源是信息和它的生产者以及信息技术的集合。

2. 信息资源的分类

按照不同的标准可以将信息资源划分为不同的种类,常见的划分标准及其类型有:

(1) 按信息资源所依附的载体分

1) 体裁信息资源。体裁信息资源是指以人体为载体并能为他人识别的信息资源。按其表述方式分为口语信息资源(谈话、授课、演讲、唱歌等)和体语信息资源(表情、手势、姿态、舞蹈等)。

2) 实物信息资源。实物信息资源是指以实物为载体的信息资源。它可分为天然实物资源和人工实物资源。

3) 文献信息资源。文献信息资源是指以文字、图形、符号、声频等方式记录在各种载体上的知识和信息资源,是目前利用最多的信息资源。



4) 网络信息资源。网络信息资源是指以网络为纽带连接起来的信息资源和以网络为主要交流、传递、存储的手段与形式的信息资源。

(2) 按信息资源传递的范围划分, 信息资源可分为公开信息资源、半公开信息资源和非公开信息资源。

(3) 按信息资源的加工程度划分, 信息资源可分为一次信息资源、二次信息资源和三次信息资源等。

3. 文献信息资源的分类

文献信息资源是信息检索的主要对象, 是本书要讨论的重点。根据不同的标准可将文献资源进行如下划分:

(1) 按文献记录载体划分

1) 书写型。书写型是指以竹简、缣帛、纸等为载体, 由人工抄写而成。

2) 印刷型。又称纸介型, 传统的文献形式, 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、方便地阅读; 缺点是存储密度低, 体积庞大, 难于保存。

3) 缩微型。缩微型是指以感光材料为载体的文献, 又可分为缩微胶卷和缩微平片。优点是体积小, 存储密度高, 便于传递、保存; 缺点是不能直接阅读, 必须借助于阅读器。

4) 视听型。又称直感型或声像型, 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。其优点是直观、形象; 缺点是无论是制作还是阅读都需借助特定设备。

5) 电子型。电子型又称计算机阅读型, 是一种最新形式的载体。它主要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, 它能存储大量情报, 可按任何形式组织这些情报, 并能以极快的速度从中取出所需的情报, 如电子图书期刊、各种联机信息库、光盘数据库、电子邮件等。其优点是存储容量大, 存储速度快, 原记录可以修改、删除或更新等; 缺点是设备投资高、价格昂贵。

(2) 按文献加工等级划分

1) 零次文献。

含义: 它是指未经过任何加工的原始文献。

作用: 不仅在原始文献的保存、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, 而且能弥补一般公开文献从信息的客观形成到公开传播之间费时多的弊病。

举例: 实验记录、手稿、原始录音、原始录像、谈话记录等。

2) 一次文献。

含义: 人们直接以自己的生产、科研、社会活动等实践经验为依据生产出来的文献, 其所记载的知识、信息比较新颖、具体、详尽。

作用: 它是科技查新工作中进行文献对比分析的主要依据。

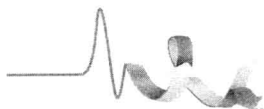
举例: 期刊论文、专利文献、科技报告、会议录、学位论文等。

3) 二次文献。

含义: 将大量分散、零乱、无序的一次文献进行整理、浓缩、提炼, 并按照一定的逻辑顺序和科学体系加以编排存储, 使之系统化, 以便于检索利用。

作用: 查新工作中检索文献利用的主要工具。

举例: 主要类型有目录、索引等, 如《中文科技资料目录》和《中国科技期刊数据库》。



4) 三次文献。

含义：它是指选用大量有关的文献，经过综合、分析、研究而编写出来的文献。通常围绕某个专题，利用二次文献检索搜集大量相关文献，对其内容进行深度加工而成。

作用：充分利用反映某一领域研究动态的综述类文献，在短时间内了解其研究历史、发展动态、水平等，以便能更准确地掌握待查项目的技术背景，把握查新点。

举例：综述、评论、评述、进展、动态等。

从零次文献、一次文献、二次文献到三次文献，是一个由分散到集中，由无序到有序，由博而精地对知识信息进行不同层次的加工过程。它们所含信息的质和量是不同的，对于改善人们的知识结构所起到的作用也不同。零次文献和一次文献是最基本的信息源，是文献信息检索和利用的主要对象；二次文献是一次文献的集中提炼和有序化，是文献信息检索的工具；三次文献是把分散的零次、一次、二次，按照专题或知识的门类进行综合分析加工而成的成果，是高度浓缩的文献信息，既是文献信息检索和利用的对象，又是检索文献信息的工具。

(3) 按出版类型划分

1) 图书。

含义：凡篇幅达到 48 页以上并构成一个书目单元的文献称为图书。

特点：带有总结性、成熟定型；出版周期长，信息传递慢；是传授知识，而不是报道最新情报。

正式出版的图书有 ISBN 代码，即国际标准书号。每一种正式出版的图书的唯一标志代码。例如：

ISBN 商品号 组号 出版者号 书名号 校验码

ISBN 978—7—118—01984—4

2) 期刊。

含义：期刊也称杂志，是指那些定期或不定期出版、汇集了多位著者论文的连续出版物。

特点：名称固定；有连续的卷、年月顺序号；出版周期短，报道速度快；数量大，内容丰富。

我国有期刊 8000 余种，多种期刊在学科、主办单位、主管部门、质量、服务等方面千差万别，尽管国家行政管理部门声明从未从行政角度对现行期刊进行级别划分，但期刊级别的观念早已深入人心。期刊等级划分通用标准如下：

① 按期刊的主管部门分级。这是期刊分级的最传统的方法，也是目前仍在使用的主要方法。按照这种分级方法，期刊被分为国家级、省部级、地市级，由代表国家科研水平的科研院所、高等学校、国家一级学会主办的学术期刊一般被认为是国家级期刊，省部级、地市级依此类推。

② 按期刊是否公开出版分级。按此标准，期刊分为公开发行人和内部发行期刊，正式出版的期刊有 ISSN 代码，即国际标准连续性出版物编号。例如：ISSN 1000—131X 为《土木工程学报》的国际标准编号。一般认为，在主办者级别相同的情况下，公开发行人期刊要高于内部发行期刊（有保密要求或非学术因素暂不便公开的除外）。

③ 以期刊质量分级。以质量分级是最具有学术价值的一种分级，主要表现为期刊获奖或入选某种目录。核心期刊的理论传入我国后，情报学界终于有了替代传统的以行政级别分级的较为科学、合理的新方法。

1992 年，《中文核心期刊要目总览》（以下简称《总览》）出版和全国优秀科技期刊评比，标志着期刊质量评价和期刊分级研究进入一个新阶段。《总览》由北京高校图书馆期刊工作研究会组织北京地区 41 所高校 200 多名人员参加编制，历时两年，从 1 万多种中文期刊中筛选



出核心期刊2156种,出版了165万字的《总览》。现在学术界认可,国家承认的分类就只有三种:一是北大中文核心(北京大学评的,简称北大核心)是北京大学图书馆联合众多学术界权威专家鉴定,目前受到了学术界的广泛认同。从影响力来讲,其等级属同类划分中较权威的一种。它是除南大核心、中国科学引文数据库(CSCD)以外学术影响力最权威的一种。

北大图书馆每四年出版一次《全国中文核心期刊要目总览》(2011年后每三年出版一次),目前最新版为2012年版。

3) 特种文献。

① 科技报告。

含义:科技报告又称研究报告和技术报告,是科学技术工作者围绕某个课题研究所取得成果的正式报告,或对某个课题研究过程中各阶段进展情况的实际记录。

特点:单独成册,所报道成果一般须经过主管部门组织有关单位鉴定,其内容专深、可靠、详尽,而且不受篇幅限制,可操作性强,报告迅速。有些报告因涉及尖端技术或国防问题等,所以一般控制发行。

② 专利文献。

含义:专利文献是指发明人或专利权人申请专利时向专利局所呈交的一份详细说明发明的目的、构成及效果的书面技术文件,经专利局审查,公开出版或授权后的文献。广义的专利文献还包括专利公报(摘要)及专利的各种检索工具。

特点:数量庞大、报道快、学科领域广阔、内容新颖,具有实用性和可靠性。

③ 会议文献。

含义:会议文献是指各种科学技术会议上所发表的论文、报告稿、讲演稿等与会议有关的文献。

特点:传播信息及时、论题集中、内容新颖、专业性强、质量较高,往往代表某一学科或专业领域内最新学术研究成果,基本上反映了该学科或专业的学术水平、研究动态和发展趋势。会议文献是科技查新中重要的信息源之一。

④ 标准文献。

含义:标准文献是公认的权威机构批准的标准化工作成果,反映了当时的技术工艺水平及技术政策。标准文献是记录人们在从事科学试验、工程设计、生产建设、商品流通、技术转让和组织管理时共同遵守的技术文件。

特点:能较全面地反映标准制定国的经济和技术政策,技术、生产及工艺水平,自然条件及资源情况等;能够提供许多其他文献不可能包含的特殊技术信息。标准文献具有严肃性、法律性、时效性和滞后性。

⑤ 政府出版物。

含义:政府出版物是指各国政府部门及其设立的专门机构发表、出版的文件,可分为行政性文件(如法令、方针政策、统计资料等)和科技文献(包括政府所属各部门的科技研究报告、科技成果公布、科普资料及技术政策文件等),其中科技文献约占30%~40%。

特点:内容可靠,与其他信息源有一定重复。借助于政府出版物,可以了解某一国家的科技政策、经济政策等,而且对于了解其科技活动、科技成果等,有一定的参考作用。

⑥ 学位论文。

含义:学位论文是高等院校和科研院所的本科生、研究生为获得学位资格(博士、硕士和学士)而撰写的学术性较强的研究论文,是在学习和研究中参考大量文献、进行科学研究的基础上而完成的。