



信息

检索与利用

王放虎 李 凯 编著

E mail



1,256,45 2,258,40 6,111,00
1,245,00 1,245,23 1,011,11

1,821,30 3,031,40 2,411,00

2,704,00 3,110,40

教育科学出版社

信息检索与利用

王放虎 李 凯 编著

教育科学出版社

内 容 简 介

为加强高职院校学生信息意识和能力的培养,本书遵循信息检索发展的历史过程及高职院校信息素质培养目标的要求,并以中文信息、现代电子网络信息为主,介绍了手工、光盘、网络检索方式的基本概念、检索原理、检索工具、信息获取渠道及信息的利用。取材新颖,内容详实、图文并茂、可读性强。

本书适合于高职院校学生作为信息检索与利用课的教科书,也可作为从事教学、科研、管理以及其他一切与信息检索相关的广大读者的参考工具书。

图书在版编目(CIP数据)

信息检索与利用/王放虎 李 凯编著

—北京:教育科学出版社,2007.1

ISBN 978-7-5041-4723-3

I. 信… II. ①王…②李… III. 信息科学-理论与实务

IV. G252.7

中国版本图书馆CIP数据核字(2007)第086912号

出版发行	教育科学出版社
社 址	北京,北三环中路46号
网 址	http://www.pep.com.cn
经 销	新华书店北京发行所
印 刷	北京国光印刷厂印装
开 本	880×1240 1/24
印 张	16.9
字 数	290千字
版 次	2007年1月第一次印刷
印 数	1-1000册
书 号	ISBN 978-7-5041-4723-3/G·262
定 价	26.00元

(版权所有 侵权必究)

前 言

信息社会,人们把信息、物质与能量一起称为人类社会赖以生存发展的三大要素。信息是促进社会经济、科学技术以及人类生活向前发展的重要因素。一个国家的科技进步和社会发展越来越取决于对信息的开发利用,谁能充分和有效地利用信息资源,谁就能抢占科学技术的制高点。

信息化的社会环境对人才的要求更高,信息素质成为现代化人才必备的基本素质之一。1984年,教育部就提出在大学生中开设“文献检索与利用”课程,“信息检索”是该课程的延续和发展。2003年,教育部教高司发布的《高职高专院校人才培养工作水平评估方案(试行)》中,把学生收集处理信息能力列入考核内容。

信息检索课的宗旨是培养学生的信息素质,使学生在增强信息意识的基础上,熟悉检索工具和系统,掌握信息检索的方法和技巧,提高自学能力,激发创新能力。目前,面向普通高校的信息检索课教材很多,而适合高职高专院校的教材较少,为此,在总结几年信息检索课教学经验的基础上,结合我院专业的特点,编写出“信息检索与利用”一书,应用于检索课教学。

本书以信息检索发展的历史过程为基本线索,以中文信息、电子网络信息为主,分为基本知识、手工检索、计算机光盘检索、网络信息检索、文献信息的利用五章。第一、三、四章由王放虎执笔编写;第二、五章由李凯执笔编写。各章初稿完成后由王放虎统改全稿并定稿。

编写过程中,为了尽力吸纳信息检索的最新资料和发展成果,我们参阅了近年来许多作者编写出版的信息检索教材、专著、研究论文以及网络上的有关信息资源;在篇章结构谋划、资料收集、内容选用等方面,还得到本院赵建民教授(馆长)、屈锡泰副教授、吴小凤副馆长及有关同志的指导与帮助,在此,我们一并表示衷心的感谢。

由于我们经验不足,业务水平有限,缺点错误在所难免,敬请各位前辈、同行、读者不吝指正。

编 者

2007年1月于杨凌职业技术学院

目 录

第一章 信息检索基本知识-----	1
第一节 几个基本概念-----	1
一、信息的概念、特征-----	1
二、知识-----	3
三、情报-----	4
四、文献-----	6
五、信息、知识、情报、文献之间的关系-----	7
第二节 信息的类型-----	8
一、按信息的表现形式划分-----	9
二、按信息的出版类型划分-----	9
三、按信息的加工层次划分-----	12
第三节 信息检索的原理-----	13
一、信息检索的概念-----	13
二、信息检索原理-----	13
三、信息检索的类型-----	14
第四节 信息检索语言-----	15
一、检索语言及其功能-----	15
二、检索语言原理-----	15
三、检索语言组成-----	15
四、检索语言的种类-----	16
第五节 信息检索方法、途径、工具、步骤-----	19

一、检索方法-----	19
二、检索途径-----	19
三、检索工具及其利用-----	20
四、检索步骤-----	21
第六节 信息检索效率评价-----	25
第二章 手工检索-----	27
第一节 中文手工检索工具介绍-----	27
一、中文图书检索工具-----	27
二、中文报刊检索工具-----	31
三、特种文献检索工具-----	37
四、参考工具书-----	54
第二节 手工检索举例-----	63
一、用分类途径检索-----	63
二、用主题途径检索-----	64
三、用著者途径检索-----	64
第三章 计算机光盘检索-----	66
第一节 计算机检索概述-----	66
一、历史和现状-----	66
二、计算机检索的特点-----	66
三、计算机检索的相关概念-----	67
四、计算机检索方法-----	68
五、计算机检索技术-----	69
第二节 联机信息检索-----	75

一、联机检索方式的发展	75
二、国际著名联机检索系统简介	75
三、我国的联机系统与数据库	76
第三节 光盘信息检索	78
一、概述	78
二、中国光盘数据库及检索简介	79
第四章 网络信息检索	117
第一节 因特网概述	117
一、什么叫因特网?	117
二、网络信息资源及其类型	117
三、因特网信息资源的特点	119
四、因特网有关技术术语	119
第二节 网络检索工具	122
一、概述	122
二、工作原理	122
三、搜索引擎分类	123
四、通用型检索工具	124
五、英文目录类搜索引擎	134
六、中文搜索引擎	138
七、中文目录搜索引擎	142
八、网络检索工具应用	146
九、检索工具对检索结果的排序	149
第三节 各类文献的网上检索	149

一、综合性文献信息检索指南	149
二、人物、机构类信息数据库	221
三、专业信息资源检索	224
四、特种文献信息资源检索	243
第五章 文献信息的利用	248
第一节 文献信息的收集、积累整理	248
一、确定文献收集的范围和类型	248
二、文献信息的收集方法	249
三、文献信息的积累	249
四、文献信息的整理	249
第二节 文献信息的选择及阅读	249
一、文献的选择	249
二、文献的阅读	250
第三节 文献信息的研究	251
一、信息研究的类型	251
二、常用研究方法简介	252
第四节 信息调研报告及其撰写	254
一、调研报告的类型	254
二、调研报告的性质	255
三、调研报告的结构	256
四、调研报告的撰写	256
参考文献	258

第一章 信息检索基本知识

第一节 几个基本概念

一、信息的概念、特征

(一) 信息的概念:

信息一词,在我国古代文献中,最早见于南唐诗人李中的《碧云集·暮春怀故人》。其诗句是:

池馆寂寥三月尽,落花重叠盖莓苔。
惜春眷恋不忍扫,感物心情无计开。
梦断美人沉信息,目穿长路倚楼台。
琅玕绣段安可得,流水浮云共不回。

诗句中的“信息”一词含义为音讯或消息,此后,历代文献多有“信息”一词出现,其基本含义是指消息、音讯。

信息作为一个科学概念,是19世纪末提出来的。由于信息的定义一直是多学科、多领域中研究的一个基本问题,故对信息含义的界定众说纷纭,据不完全统计,信息的定义有100多种,至今无法统一。其原因:

第一,信息本身的复杂性,它是一个多元化,多层次,多功能的综合物;

第二,信息科学是一门新兴学科,它的许多分支学科仍在随着社会、经济 and 科技的发展而发展;

第三,人们出于不同的研究和使用目的,从不同的角度或层次出发,对信息概念就会作出不同的解释。

故一些学者建议将信息的要领分为三个层面来把握:

1.本体论意义上信息概念

本体论(即哲学)意义的信息是指“事物运动的状态与方式”。“事物”泛指一切可能研究的对象,包括外部世界的物质客体,也包括主观世界的精神现象。“运动”泛指一切意义上的变化,包括机械运动、物理运动、化学运动、生物运动、思维运动等。运动方式是指事物运动在时间上所呈现的过程和规律。“运动状态”则是事物运动在空间上所展示的态势。

2.各学科领域中的信息概念

各学科领域中的信息主要指进入人认识领域和传播领域,可以被理解或被

接受的消息、情报、知识事实、数据等。这一定义有以下要点：其一，信息是指已被人类认识、理解、开发利用的信息。如 N.维纳认为：“信息是我们适应外部世界并且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。”“生物以及具有自动控制系统的机械系统，通过感觉器官和外界交换的一切内容。”其二，信息是认识过程中不确定性的消除或减少。如申农（又译作香农）认为：“信息是能够用来消除不确定性的东西”。其三，信息不是已知消息的重复，而是未知的内容。

3. 大众交流信息

主要指在日常生活、学习和工作中广泛使用的、在各种媒体中普遍通用的“信息”。这一概念是将相互之间交流的所有内容都称之为信息，是被理解、被接受、时时、处处都存在的信息，是不言自明的信息，诸如经济信息、政治信息、文化信息、工业信息、农业信息、电子信息、生物信息、计算机信息、广告信息等。大众交流信息是最大众化、最广泛和最通用的信息概念。

中国国家标准 GB4894—85 关于信息的定义：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事件的不定性减少。

（二）信息的特征

1. 依附性

信息的记载、存储、传播与交流，须依附于某种载体，以某种载体的形式表现出来，如语言文字、声音图像、纸张胶卷、磁带光盘、人脑等等，都是信息的载体，没有载体就没有信息。

2. 可加工性

信息数量庞大，质量高低不一，信息价值的发挥需进行加工处理。原始信息可加工成二次信息，二次信息通过分析、研究综合，又可加工成三次信息，每次加工，都可改变原有信息的结构，并赋予新的价值。

3. 传递性

信息在运动中产生，在传递中发挥价值。从时间上，信息可从某一时间纵向地传递到另一时间，也可以从某一地点横向地传递到另一地点。信息的获取利用以及信息的反馈，必须借助于信息的传递。

4. 广泛性

信息广泛存在于事物中，世界万物的运动，人间万象的更迭，都蕴含有信息。当今社会，无论科学研究、金融商贸，还是政府决策、民间活动，都与信息密切相关。信息广泛关系到人们工作生活的各个方面。

5.共享性

信息价值的实现,需通过对信息的使用。与实物不同,同一信息可以同时被两个以上的多个用户使用,而且,并不因为信息的多人、多次重复使用而丢失其内容。

6.时效性

信息的价值与时间有关。客观事物随着时间的流失而不断变化,在某一时间阶段产生的反映这一客观事物的信息,在价值上也会发生变化。一方面,一些信息失去了原有的价值,如股票信息;另一方面,一些信息会变得价值百倍,如古籍文献,名人字画。

(三) 信息源

1.信息源的概念

是指人们为满足信息需求而借以获得信息的一切来源。因此,凡是能产生、存储、加工、传递信息的社会活动场所、机构、人物、产品和自然物质都是信息源。

2.信息源的类型

信息是无形的,它是靠以语言符号、声音或图像等形式存储在各种介质的载体上来体现的。所以信息源可分为文献信息源和非文献信息源。文献信息源是当今记录信息和我们检索所需信息的主体,后面我们将详细介绍。非文献信息源是非记录性信息源,由于其不易积累、存储和保管,过去不列为信息管理对象,但其具有传递信息直接方便、迅速及时、针对性强等优点,日益受到人们的重视。

非文献信息源分口传信息源、实物信息源两类:口传信息源是存在于人脑,借助口耳相传的一种信息源。其传递信息的方式如交谈、讨论、听报告、听广播、看电视、看电影等。口传信息获取的主要目标首先是权威人物、专家、教授、企业家等。实物信息源它是存在于自然界和人工制品中,借助于物体表达的一种信息源。其传递信息的方式有实践、试验、参观、采集标本、举办展览等。实物信息获取的主要目标首先是高技术实物,其次是实物样品。

二、知识

知识是人们在改造客观世界的实践中所获得的认识和经验的总结。是人的主观世界对客观世界的概括和如实反映。知识是大量有组织的信息,是关于事实和思想的有组织的陈述。提供某种经过思考的判断和某种实验的结果。

根据经济合作与发展组织(OECD)出版的《以知识为基础的经济》报告,知识可分为如下四类。

第一类 “知事 (Know-what)”, 指关于事实方面的知识, 也可理解为 Know-when、Know-where; 即在什么样的时间 (Know-when)、什么样的地点或条件下 (Know-where) 能解决什么样的问题;

第二类 “知因 (Know-why)”, 指自然原理和规律方面的科学理论, 知识的生产是在专门研究机构如实验室和大学完成的;

第三类 “知道怎样做的知识 (Know-how)”, 指做某些事情的技艺和能力, 被称为技术情报和商业秘密, 其典型是企业开发和保存于其内部的技术诀窍或专有技术;

第四类 “谁以及是怎样创造知识的 (Know-who)” 侧重创造思想、方法、手段、过程以及特点等的了解。

知识的来源

第一类是直接源于产生信息的客观事物; 第二类是通过信息载体或媒介 (文献、电视、广播、他人等) 的传递、交流而间接获得。

三、情报

(一) 情报的定义

情报的概念也是一个发展中的概念。其定义有十多种, 如有人说“情报是知识的传递并起作用的部分”。英国 B.C 布鲁克斯认为“情报是使人类原有的知识结构发生变化的那一部分知识”。钱学森说: “情报是为了解决一个特定的问题所需要的知识”。

(二) 情报的基本属性

1. 情报的知识性

知识性是情报的基本属性之一。一般地讲, 情报是一种新知识。科技的发展意味着新的知识的产生和陈旧知识的更替。创造发明、科研成果、新技术、新工艺、新设计、新产品、新理论、新事实、新决策等, 都是新知识。这些新知识一般均可视为情报。但称为情报的知识并不排斥时间上的“不新”。就是说, 凡原先不知道的, 而现在又迫切需要的知识, 都属于“新”的知识。所以情报以知识为实体, 情报的本质就是知识。知识是情报的基本属性。

2. 情报的传递性

情报的本质是知识, 而这些知识如不进行传递, 仍不能称其为情报。如一个人的知识无论怎样渊博与新奇, 如果不记录、不传递、不交流, 都不能称为情报。所以传递性是情报的又一基本属性。

3. 情报的效用性

情报是一种有价值、有效用的知识。其效用性在于:

(1)能启迪思路、开阔眼界、增进知识。

(2)能执行决策、实施规划、改变面貌。

(3)能提高人类改造客观世界的能力。情报的效用价值,具有客观鉴定标准,它取决于其所含知识的可靠性、新颖性和时间性等因素。情报均须“有用”,当用户不需要时,百分之百的知识也不能称为情报。我们一方面强调“情报是普遍存在的社会现象”,另一方面又强调“情报的专指性”,对一个具体用户来讲,其所需情报并不是随手可得,而是要付出艰苦的劳动,才能取得有价值的情报。

概括上述内容情报定义可表述为:情报是特定时间、特定状态下、给特定的人提供的有用的知识。

(三) 竞争情报

竞争情报是市场竞争过程中产生的。关于竞争情报的定义,目前尚无定论,总体上国外行业人士倾向于这样一种解释:“所谓的竞争情报:是关于企业竞争对手的能力、弱点和意图的信息,是关于竞争环境、竞争对手和竞争策略的情报研究,是适应市场竞争的需要,赢得自身的竞争而出现的一种新的情报概念。竞争情报一般是指竞争主体为保持竞争优势所需要的一切有关竞争对手和竞争环境的情报,广义上也含相关的情报活动。

国内学者从广义角度对竞争情报定义为:是指一个国家、组织、团体为了在竞争中取得优势所需要的和进行收集、整理、加工、分析、研究的所有信息。从狭义上的定义,很大程度上是从经济信息方面考虑的企业竞争情报,也就是某企业为了在市场竞争中赢得有利地位而收集和处理的、与自身生存和发展密切相关的各种信息,主要是竞争对手和企业本身的经济方面的信息,也包括与企业经济发展的政治、法律、社会文化等环境信息。

竞争情报的作用在于:它有助于发展企业的核心竞争能力,是强化和改变企业发展战略的重要基础。竞争情报的作用主要在于以下三大作用:充当企业的预警系统、充当企业的决策支持系统、提高企业的竞争力、从细处来说竞争情报的作用在于:生产情报产品;可以支持企业实施高效管理;可以帮助企业领导实现科学决策;可以引导企业创新的正确方向;可以促进企业的文化建设。

搜集对方竞争情报的方法概括起来有两种:

第一种,采用一般方法:

(1)通过技术、销售和供应等部门日常的业务渠道收集、捕捉信息,特别是营销人员是获得市场信息的重要途径。

(2)积极参加各种相关的展览会。收集国内外新产品、新技术、新工艺、新

设备以及市场动态等最新情报。

(3)定期去各类图书馆、情报信息研究所以及专利局查阅所需的科技文献、专业论文以及专利文献,可获得准确可靠的情报资料。

(4)关注各种专业报刊杂志,从中可获得很多有价值的信息。

(5)积极参加相关协会、学会举办的学术研讨会、交流会。通过学术交流、与会者的发言、提供的资料文件以及相关间的讨论和质疑可获取在一般场合下难以得到的信息情报。

(6)通过和国内外大专院校、科研院所以及企业集团建立和保持良好的业务合作关系、在相互技术合作和商务往来中可以了解有价值的信息情报,同时可提高企业的知名度。

(7)广交朋友、建立良好的人际关系网,特别是产品经销商、原料及设备供应商这一渠道,有了良好的人际关系往往可以从这一渠道了解到竞争对手内部鲜为人知的情报。

(8)聘请业内专家和顾问,有利于获得高层次、高质量的信息情报。

(9)组织专题调研小组进行可行性调研。

(10)利用因特网广泛搜索国内外科技动态、最新科研成果、实用技术、产品信息、经贸信息和市场行情等相关情报。

第二种,采取特殊方法:

(1)利用可能的参观和考察机会深入到被调查方了解所要的情报。

(2)采用委托咨询方式进行相关内容的调研。在已有信息情报不足以进行决策时,可以委托专业咨询公司进行专题研讨,这样得来的情报资料相对比较客观、公正可靠,可以帮助企业决策者摆脱主观因素,以比较冷静的思维对竞争形势进行判断和决策,减少因策略失误给企业带来的损失。

(3)利用反证工程进行调研。可通过合法手段取得竞争者的产品实物和样品,使用先进的科技手段对其进行解剖和深入的分析,了解其结构特点、化学成分、使用功能以及制造工艺手段和水平,这对改造自己产品的性能、提高产品的竞争力具有十分现实的意义。

四、文献

(一)文献的定义

“文献”一词的出现,在我国已有两千年的历史。在古代,文献的含义有两方面:一是文章典籍,即书本记载;二是先贤及其口头议论。

纸张载体的出现,以印刷技术为记录方式的文献功能扩大,文献第二方面的含义弱化,文献含义为“记录有知识的一切载体”。

在国外,文献含义的说法有:“一定范围的文件的总和”、“某一具体领域的全部资料的总汇”、“可作为一个单元处理的,在载体内,载体上或依附载体而存贮有信息或数据的载体”。等等。

目前有关文献较权威的定义主要有两个:

一是《文献情报术语国际标准(草案)》(ISO / DIS5127)的定义,“为了把人类知识传播开来和继承下去,人们用文字、图形、符号、声频、视频等手段将其记录下来,或写在纸上,或晒在蓝图上,或摄制在感光片上,或录到唱片上,或存贮在磁盘上。这种附着在各种载体上的记录统称为文献。”

另一个是中国的国家标准。是各种媒介和形式的信息集合,包括文字、声像印刷品、电子信息、数据库等。

(二) 文献的构成要素

1. 文献信息是文献的内容。

2. 符号系统(指图画、文字、公式、图表、编码、声像和电磁信息等)是信息的携带者。

3. 载体是符号赖以依附的“寄主”。随着生产力的发展,存贮和表达人们思想的物质载体从金石、竹简、羊皮、丝帛、纸张发展到用感光介质和磁性介质。从非人工材质的泥版、岩石、石板、兽骨、木板、竹片、兽皮、树叶、桦树皮等到各种人工材质,如无机材质(陶、砖瓦、瓷,玻璃等)、金属材质(青铜、铁、铝、金银等)、高分子材质(帛、纸、胶片、醋酸纤维等)、复合材质(磁带、光盘等)。纸质文献已经多得便于快速传输信息、高效查阅和高密度存贮了,于是其他介质的文献应运而生。但是,由于纸质文献具有价格低廉、质地柔软、易于书写、携带和收藏等其它一些载体所无法比拟的性能而成为文献家族的主干。

4. 记录方式是将文献的符号进入载体的方法和过程。按记录方法分可有手工记录、机械记录、光记录、电记录、声记录 and 磁记录。其中最常见的是印刷。随着科技的发展,文字记录可以转变为数据记录,并用电子方法存贮到磁介质上去。此外,还有光学字符识别(OCR)技术,这种崭新的技术,它通过光学方法对字符、标记表示的书面数据进行自动识别,转换成机器可以处理的信息,这样就实现了纸质文献信息向电子信息的转换,而且可以通过不同接口装置向不同设备输入数据,这些设备进而与通信网络相连,实现远距离数据传送。

五、信息、知识、情报、文献之间的关系

信息、知识、情报、文献之间的关系:知识来源于信息,是理性化、优化和系统化了的的信息;情报是解决特定问题的知识和智慧,是被激活的那部分知

识；文献是它们的载体。在一定条件下，它们之间可以互相转化。信息、知识、情报、文献之间的关系可用下面图例来说明：

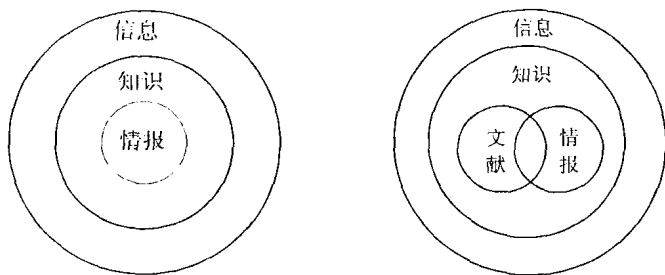


图 1-1 信息、知识、情报、文献关系图

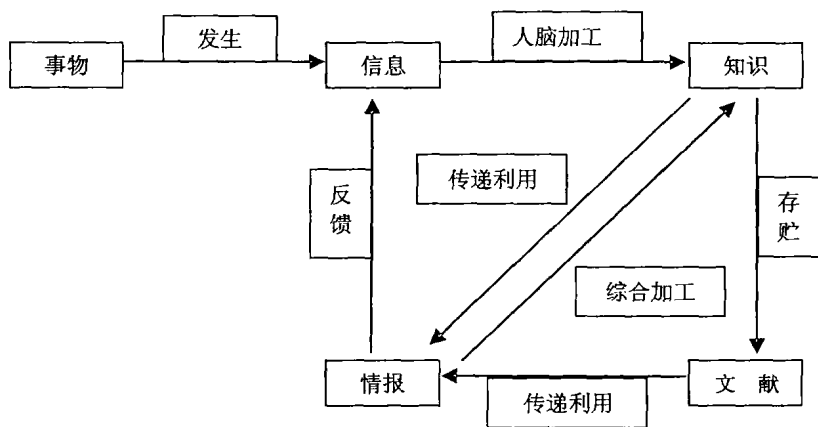


图 1-2 信息、知识、情报、文献相互转换关系图

第二节 信息的类型

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，所以很难用统一的标准进行分类。以下从信息的表现形式、出版类型和加工层次三方面加以介绍。

一、按信息的表现形式划分

(一) 文字信息

文字是人们为实现信息交流、通信联系所创造的一种约定的形象符号。广义的文字还包括各种编码,如电报编码、二进制编码。信息的内容表现于这些文字、符号、代码的结构属性之中。如二进制码“0”和“1”的不同排列,就代表着不同的信息内容。

(二) 图象信息

图象(形)是一种视觉信息,它比文本信息直接,易于理解。如一幅画、一张相片、一部电影。大自然的客观景象、人工旅游景观可视为抽象或间接的图象信息。

(三) 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、测量数据等。广义的“数据”特指在计算机网络中存储、处理、传输的二进制数字字符编码。

(四) 语音信息

人讲话实际上是大脑的某种编码形式的信息转换成语言信息的输出,这是一种最普通的信息表现形式。

音乐也是一种信息形式,是一种特殊的声音信息,它是通过演奏的方式表达丰富多彩的信息内容的。

二、按信息的出版类型划分

(一) 图书

是以印刷方式单本刊行的,内容比较成熟、资料比较系统、有完整定型的装帧形式的出版物。包括专著、教科书、科普读物、专业参考工具书等。图书的特点是:内容比较系统、全面、成熟、可靠,是人们学习、研究不可缺少的信息来源。但出版周期较长,体积大,更新速度慢,传递信息速度慢。图书重要的特征之一是国际标准书号(ISBN)。

(二) 期刊(杂志)

期刊是指名称固定、开本一致,汇集了多位著者论文,定期或不定期出版的连续出版物。期刊内容新颖,报道速度快,信息含量大,是传递科技情报、交流学术思想最基本的文献形式。期刊情报约占整个情报源的60%~70%,它与专利文献、科技图书三者被视为科技文献的三大支柱,也是检索工作中利用率最高的文献源。大多数检索工具也以期刊论文作为报道的主要对象。对某一