

文献信息检索新论

WENXIAN XINXI JIANSUO XINLUN

◎ 廖连生 编著



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

文献信息检索新论

廖连生 编 著



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

图书在版编目(CIP)数据

文献信息检索新论/廖连生编著. —杭州:浙江大学出版社, 2014.1

ISBN 978-7-308-12864-3

I. ①文… II. ①廖… III. ①情报检索 IV. ①G252.7

中国版本图书馆CIP数据核字(2014)第019770号

文献信息检索新论

廖连生 编著

责任编辑 邹小宁

文字编辑 李春宝

封面设计 王聪聪

出 版 浙江大学出版社

(杭州市天目山路148号 邮政编码 310007)

(网址: <http://www.zjupress.com>)

排 版 杭州教联文化发展有限公司

印 刷 浙江云广印业有限公司

开 本 787mm×1092mm 1/16

印 张 12.25

字 数 283千

版 次 2014年1月第1版 2014年1月第1次印刷

书 号 ISBN 978-7-308-12864-3

定 价 31.20元

版权所有 翻印必究 印装差错 负责调换

前 言

文献信息检索作为一门学科渊源久远,可以追溯到20世纪中期。从世界范围来看,文献信息检索作为一门学科发展,主要可以分为两个阶段。

第一阶段主要在20世纪80年以前,处于计算机还未普及,互联网还未兴起之前,那时候文献信息检索技术的研究和试验,都是采用人工标引的方法,在一个人工或者说是虚拟的环境中进行的,都是基于以“提问—检索”模式为核心的相关理论和方法为基础的,传统的手工标引、规范化的检索语言、线性的书目数据库、预定的检索策略以及以回答检索提问为主的被动服务方式是该时期文献信息检索的主要概念体系、方法和原则。第二阶段是在20世纪90年代至今,计算机技术和Internet网络技术的迅猛发展,信息资源的内容更为广泛,涉及各个学科和人类生活的方方面面;信息资源的类型也发生重大变化,除了传统的纸质信息外,音频、视频、磁质、网络信息占据越来越重要的地位;网络化与数字化技术将分布在世界各地的信息资源和数据库联为一体,为人们提供了快速、有效地跨越时空和地域的国际化交流的信息平台,这就使得传统相对集中的、相对规范的、人工环境的检索模式和系统面临着挑战。所以网络信息检索逐渐成为文献信息检索的主导,“浏览—查询”模式逐渐替代“提问—检索”模式成为网络时代信息检索的主导模式。特别是近几年来,云计算技术在计算机、网络和图书馆领域的出现和发展,使传统的文献信息检索出现了更为新颖的概念、技术、理论和学术热点,值得学术界重视和研讨。

正是由于计算机技术、网络技术和图书情报学科的不断变化,文献信息检索领域出现了一些新知识、新技术、新概念和新的学术讨论热点,笔者觉得很有必要对传统的文献信息检索内容进行一些必要的、力所能及的更新和补充,于是就产生了编著此书的念想,也是把书名定为“新论”的缘由。

本书系统、全面、深入和详尽地阐述了文献信息检索的传统概念、技术和理论。包括信息、信息检索和信息检索工具的详尽阐述,还有各种不同类型的信息资源载体如图书、报纸、期刊、特种文献、网络文献的基本知识和检索技术的系统而详尽阐述。更为重要的是,笔者在此基础上根据网络时代信息技术的发展,对文献信息检索领域中出现的新概念、新技术、新论点、新理论和新热点都进行系统地阐述,是对传统的文献信息检索学科知识的进一步更新和延伸。本书的编著都尽力做到内容的继承性和发展性相统一,知识的重点性和广泛性相结合,价值的实用性和持续性相辉映。

在本书的编著过程中,得到了许多部门、师长和同行的大力支持和帮助,在此笔者

一并致以衷心的感谢!在本书编著过程中,参阅了大量的相关专著、教材、论文和网页,引用了部分论点,限于篇幅,未能一一列出,仅列出主要的参考文献资料,在此,特向有关作者致歉并表示诚挚的谢意。

由于信息资源、信息技术的日新月异,以及互联网的迅猛发展,文献信息检索技术更是不断地变化、更新和完善,加之笔者水平和精力有限,书中难免存在遗漏和不妥之处,敬请读者批评和谅解,如能提出宝贵的意见和建议,以便今后不断地改进和完善,更是不胜感激。能有百家之言,学术方能百花齐放,繁荣昌盛!

著 者

2013年10月15日

目 录

第一章 文献信息检索基础概述及新论	1
第一节 信息基本知识概述及新论	1
第二节 文献信息基本知识概述及新论	7
第三节 文献信息的揭示、组织与特征描述概述及新论	17
第四节 文献信息检索概述及新论	19
第二章 文献信息检索工具基础概述及新论	23
第一节 文献信息检索工具概述及新论	23
第二节 检索工具书概述及新论	25
第三节 参考工具书概述及新论	30
第三章 文献信息检索工具常用排检法概述及新论	38
第一节 字顺排检法概述及新论	38
第二节 类序排检法概述及新论	41
第三节 其他排检法概述及新论	43
第四章 图书文献信息检索方法概述及新论	46
第一节 图书基本知识概述及新论	46
第二节 综合性图书的检索概述及新论	51
第三节 学科性图书的检索概述及新论	58
第四节 电子图书的检索概述及新论	63
第五章 期刊文献信息检索方法概述及新论	67
第一节 期刊的基础知识及新论	67
第二节 期刊信息资料检索概述及新论	78
第三节 常用数字期刊数据库检索概述及新论	85
第六章 报纸文献信息检索方法概述及新论	95
第一节 报纸的基础知识概述及新论	95
第二节 报纸信息资料检索概述及新论	98



第三节 网络读报软件简介概述及新论	102
第七章 特种文献信息检索方法概述及新论	105
第一节 科技报告检索概述及新论	105
第二节 专利文献检索概述及新论	114
第三节 标准文献检索概述及新论	127
第四节 会议文献检索概述及新论	137
第五节 学位论文检索概述及新论	141
第八章 网络信息资源及搜索引擎检索	147
第一节 网络信息资源概述	147
第二节 网络信息资源检索	151
第三节 网络信息检索工具	155
第四节 互联网时代的世界图书馆	158
第九章 文献信息资源的积累与合理使用	163
第一节 文献信息资源的搜集与积累	163
第二节 文献信息资源的合理使用	168
第三节 学术规范	175
附 录	182
高等学校哲学社会科学研究学术规范(试行)	182
关于加强学术道德建设的若干意见	184
参考文献	188

第一章 文献信息检索基础概述及新论

第一节 信息基本知识概述及新论

一、信息的概念

“信息(Information)”一词是1948年“信息论”的奠基人美国人克劳德·香农(C. E. Shannon)提出来的,他在研究信息理论时认为“信息是一种消息”。控制论的创始人美国人维纳(N. Wiener)在他所著的《控制论和社会》一书中认为,“信息是人们适应外部世界并且使这种适应反作用于外部世界的过程中,同外部世界交换的名称。”

在我国,“信息”一词早有记载,最早出自南唐诗人李中的诗词《暮春怀故人》:“梦断美人沉信息,目穿长路倚楼台”,这里所说的信息是指“音讯、消息”的意思。信息一词在古代使用频率也挺高,但是一直没有一个统一的标准,以前人们对信息的理解主要是指消息或者某种情况的意思。汉语“信息”在《尔雅·释诂》中作“讯”解释;在《辞海》中作“音讯、消息”解释。我国国家标准《情报与文献工作词汇基本术语》(GB 4894—85)给信息下的定义为“信息是物质存在的一种方式、形式或者运动状态,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,可以使消息所描述的不确定性减少。”

当今社会是信息社会,信息是一个极为广泛的概念,是生物信息、自然信息、人类信息等多种信息的总称,它普遍存在于自然界、生物界和人类社会之中。围绕信息而出现的资源、信息技术、信息产业、信息化社会和社会信息化等术语不胜枚举。信息的观点、概念和方法已经被政治、经济、科技、文化、生产等各个领域所接受和应用。我们可以从多个角度去揭示信息的多重属性。

从自然界的角度看,客观事物的普遍属性通常可用它的运动性、时空性、能量、反映性、质料、系统性等表示。其普遍性既是事物千差万别的表现,又是事物之间相互联系的内容。信息概念的引入旨在说明信息既不是物质,也不是能量,而是依附于自然界客观事物而存在的,即只要物质存在,就有表征其特性的信息存在。例如:“一叶而知秋”,等等。可以认为,信息实际上是客观事物运动状态、时空特性、能量大小、质地、材料、系统特征、相互联系等一切反映事物客观属性的总称。信息具有比客观事物更强的一般性和普遍性。

从人的主观认识角度看,信息是存储于人脑中的思想、观念、知识等形态。它既是

每个具体人这一特殊事物属性的一种表征,又是外部客观事物属性在人脑中留下的印记,是物质反映属性的高级形式。哲学家把人的这种直接接受客观事物信息的功能称为人的自然信息功能。人在自然界的活动是不断排除信息传输过程中的噪声和干扰,选择、过滤和提取正确信息的过程,或者说,是通过感觉、直觉、知觉、判断等阶段对来自不同方位的信息综合转换,使之成为对人自身活动有益的指导和补充的一种存在。

从人类技术和文化的角度看,信息体现在一切人为的技术和有关技术有关的事物特性之中。信息的存储在于它以某种编码的形式存在于某种介质中,如书本上的文字信息,光盘上的录音信息,照片上的图像信息,计算机中的数据信息等。这些信息既是这些人造事物的表征,也是人的知识、文化、艺术和技术水平等属性的反映。

二、信息的特征

信息的定义揭示信息的本质属性,信息来源于物质,而又不是物质本身;信息来源于精神世界,但也不限于精神领域;信息最终归结于物质的普遍性,是物质的运动状态与运动方式。信息的物质属性决定信息的一般属性特征,信息有属性特征如下。

1. 客观性

任何物质都是信息的母体,信息是对物质存在的状态、性质与特征变化的客观反映,而这种反映物质状态、性质与特征的信息是客观存在的,这种客观存在的信息可以被人类感知、获取、整理、传递和利用。

2. 普遍性

信息既不是物质,也不是能量,是客观事物普遍性的表征。不管是自然界还是人类社会,每时每刻都在产生信息。只要有物质存在,就有物质运动,有物质运动就有信息产生,信息产生是无限的,这就决定信息无处不在,无时不有,物质运动的普遍性决定信息的普遍存在性。

3. 依附性

信息本身是看不见、摸不着的,也不能独立存在和自动交流。信息只有借助某一载体或者媒介才能被交流和利用、继承和传播共享,信息不可能存在于任何物质之外。信息可以通过文字、语言、符号、代码等方式依附在纸张、胶片、磁盘、光盘等不同的载体上,其内容不因为记录方式或载体不同而发生改变。信息因依附这些载体而变成人类的一种广泛的信息资源与财富。

4. 时效性

物质世界中事物是不断变化的,而信息则跟随事物的变化不断变化,无论是针对信息整体还是某条具体的信息,它的信息内容都会随着时间的变化而变化。当人们将特定时间的事物运动的信息通过输入、处理、输出等提取过程时,由于信息的提取与信息本身状态的时间差和事物运动规律导致信息的效用逐步减弱,以至失去价值。信息价值的体现一般分为四个阶段:升值期,峰值期,减值期和负值期。信息在不同周期呈现不同的价值,在现代信息社会,信息的应用周期越来越短,如不及时把握与利用最新信

息,信息就会贬值,只有充分重视和发挥信息的时效性,才能将信息转化为时间和金钱。这表明信息的时间价值与信息经济价值,即所谓信息的时效性。

5. 可传递性

信息的价值主要体现在发挥和利用上,而信息的发挥和利用则体现在信息的传递上。信息的传递可将信息从一个时段传递到另一个时段,从一个地方传递到另外一个地方。信息的传递分为时间传递和空间传递,信息的存储属于时间上的传递活动,而空间上的传递活动就是我们通常所说的通信。因为信息具有可传递性才使得人们可以共享信息,在共享的基础上才能使信息发挥更大的作用。

6. 可存储性

信息是具体存在的,它不是可以任意想象或者随意创造的物质,它的可知性使信息可以通过记忆、收集、整序、综合、加工等方式存储,并不断地积累丰富。大脑对信息的存储就是我们通常所说的记忆。人类对信息的存储主要表现在载体上,诸如图书、期刊、档案、资料、光盘、磁盘、磁带等介质存储大量的信息,信息通过有序存储可以实现信息的积累,以提高其质量和使用价值。人类对信息的存储和积累是为了以后能更加方便、快捷、有效地利用信息。

7. 可转换性

人类为了更好地对信息转换、加工和传递,发明创造一系列的相关技术、方法和设备,运用这些技术、方法和设备对信息编码、压缩和有序化转换等,使信息以不同的形态和方式来表达和转换。例如,同一信息既可以用语言文字来表达,也可以用图像、声音来表达;同类的信息也可以依附于纸张、胶片、光盘、磁盘等不同的物质载体,而这些物质载体之间也可以相互的转换,并且其信息内容保持不变。信息的可转换性也称为信息的有序性,对信息一系列的加工整理,使无序信息变成有序信息,有序信息便于流通、传递和利用,同时有利于信息继承和再生。

8. 可伪性

信息的可伪性包括两个方面,一方面,由于种种原因,人类借助仪器或者载体使人类在感知、接收和识别物质信息过程中产生对客观事物的错误认识和判断;另一方面,信息的相对独立性使信息脱离原事物后,人们凭主观认识来理解信息,由于人们的认识能力或动机不同,就会存在对信息的片面认识,产生虚假信息,即伪信息。

9. 动态不定性

信息是动态不定的,信息所反映出的总是在特定时刻事物运动的状态和方式,当人们将该时刻的信息提取之后,事物仍然在不停地运动。古希腊哲学家亚里士多德曾言:“人类不能两次踏进同一条河流”,这是信息不定性的经典的形象写照。

10. 差异性

信息对于不同的用户的使用存在差异性,由于人们对信息的需求、理解和判断力有别,所以相同的信息对不同的使用者有着不同的作用,体现不同的价值。

11. 可预测性

人们利用成熟的理论和相关的仪器设备,在掌握大量信息的基础上,通过有计划、有目的地预测、推理等方法来获得未来的发展,未来的发展信息一般超前于现实信息,

同时我们又称信息的超前性。

12. 共享性

信息与物质和能量的本质区别是信息具有物质和能量不具备的共享性。一般而言,物质和能量归所有者拥有,在互相交换中所有权也随着交换,一方失去,一方获得。但信息不同,同一内容的信息可在同一时间不同地点由两个或者多个用户获得使用,并且在信息的传递与交换过程中,原信息提供者并不会因为提供信息而失去原有的信息内容和信息量。换言之,就是不会因为一个人得到了信息,而另一个人就失去了信息。信息可以无限制地复制、传播和重复使用,可以分配给大家所共享,信息不会因为多人重复复制和使用而减少。相反,信息的广泛交流可能产生更多的信息。英国文学家萧伯纳说得好:倘若你有一个苹果,我有一个苹果,我们彼此交换,我们双方仍然各有一个苹果;但是如果你有一种思想,我也有一种思想,我们彼此交流这些思想,我们两个人就各有两种思想。所以说,信息的共享性使信息能重复地使用,可以大大地提高信息的利用率和使用价值。

三、信息的分类

信息广泛地存在于客观世界,信息是一个大系统,与人类智能活动有关的知识、技术、科学、文化、社会等紧密地联系在一起,我们要对信息全面系统的认识,就必须对信息分类,但是信息涉及的范围如此广泛,存在的形式又多种多样,以至于很难用统一的标准来分类。所以,我们必须根据人们对信息的理解、判断和需求的不同,从不同的角度按不同的标准对信息分类,这样才有助于人们全面地了解与认识信息,并促进我们更好地利用信息。

1. 按信息的表现形式划分

(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。广义的文字还包括各种编码,如 ASCII 码、国际电报与单元代码、计算机的二进制数字编码等等。这些文字、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中。如基本笔画的不同组合,字和字母的不同组合,二进制码“0”和“1”的不同排列等,都表示不同的信息内容和含义。

(2) 图像信息

图像信息是以图像、图形为主要表现形式的一种视觉信息,图像信息较之于文字信息更直观,更易于理解。它包括人工图像信息和自然景象信息,人工图像信息是指人类通过某种制图软件和工具制作的图形和图像等从视觉上所反映出来的信息;自然景象信息是指自然界中客观表现出的某种间接或是抽象的图形图像信息。

(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、测量数据、气象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中网络中存储、处理、传输的二进制数字字符编码,文字信息、图像信息、语言信息已经从自然界中直接采集的各种自然信息均可转换为二进制代

码,网络中的数据通信、数据处理和数据库等是广义的数值数据信息。

(4) 语音信息

语音信息主要是以声音为表现形式,包括人们日常的交流、演讲、吹拉弹唱等所产生的信息。人的日常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,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,它是通过演奏方式表达丰富多彩的信息内容的。

2. 按信息的出版类型划分

(1) 图书

包括专著、教科书、科普读物及专业参考工具书等。图书是经编著者精心选择、反复斟酌后写成的,其内容系统、成熟、定型,信息经过筛选,可靠性强,是人们从事学习、研究不可缺少的信息来源。

(2) 期刊

是指按同一专业领域定期或不定期出版的连续出版物。它出版数量大、周期短、内容新颖,能迅速反映国内外的各种学科专业的水平和动向,是人们获取一般基础理论研究知识的重要信息来源。期刊按内容分,有综合性的和专业性的;按性质分,有学术性的、技术性的、消息性的、检索性的以及通报性的;按对原文的压缩程度分,有目录、索引、文献、快报、文献指南、书目之书目等。

(3) 政府出版物

是指各国各级政府及其附属机构出版的文献资料信息,它主要包括社会科学和自然科学两大类。

(4) 科技报告

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。科技报告保密性强,难以获得。

(5) 专利文献

是指发明人向政府部门(专利局)递交的、说明自己的创造发明的技术文件,同时也是发明所有权的法律性文件。专利文献包括专利说明书、专利公报(摘要)、商标、设计公报及其检索专利的工具等。专利文献具有技术性、新颖性、独创性、实用性等特征,是重要的技术经济情报来源。

(6) 会议文献

是指国内外学术团体举行的专业会议上发表的论文与报告。与期刊相比,会议论文具有传播情报信息更迅速的功能。它反映某学科、专业的最新成果和发展水平动向,是科研工作中不可缺少的情报来源。

(7) 学位论文

是指高等院校和科研机构的毕业生、研究生(硕士或博士)攻读学位而撰写的毕业论文。它经专家评审、鉴定通过,一般说具有学术性强的特点,往往有独到的见解。

(8) 技术标准和规范

主要是指技术规范、技术标准、操作规范、建议、准则、术语、专门名词等在内的各种

技术文件。在标准实践领域中,技术标准和规范在适用范围方面是有区别的。前者是一种得到管理机构认可,适用于一定专业领域的技术规范,具有法定性;后者是指对产品、材料、工艺流程或技术特点的说明书,它仅以满足买方或工业规定的要求为准则。技术标准主要包括尺寸标准、材料标准、性能标准、方法标准、操作规程、术语和图形符号标准、文献标准等。

(9)产品样本说明

是指制造厂家和产品销售者介绍其产品的宣传性出版物。它介绍的是已投产或行销的产品。通过产品样本说明书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情报。由于产品样本说明书附有大量图表、产品特征曲线、方程等,因此具有直观的特点。同时,厂家为了推销产品,往往免费赠送,使产品样本说明具有易于获取的优点。产品样本说明书除了直接出版发行之外,还常常被包含在一些贸易刊物、企业介绍、数据手册中。

(10)技术档案

是指在科技生产活动中形成的一系列以工程技术图纸、任务书、协议、合同、设计方案以及与此有关的调查统计数据等材料组成的文件。技术档案具有技术性、适用性、保密性等特征。

3. 按信息的加工程度划分

(1)一次信息

一次信息是人们研究或创造性活动成果的直接记录,一般指公共出版的图书、期刊论文、科技报告、会议文献、学位论文、发明专利等。不管其新信息存储于何种物质载体及出版的版次,只要是原始资料就是一次信息。一次信息零碎、分散、无序,有点很难获取。

(2)二次信息

二次信息是对一次信息加工、整理而成的。如目录、文献、索引等各种数目数据库是二次信息的核心。二次信息的形成是信息从分散、无序到集中、有序化的数目控制过程。

(3)三次信息

三次信息是对一、二次信息综合、分析等深加工的产物,如评论、进展报告、述评、年鉴、指南、期刊书目等。

信息还有其他多种分类方法:

按哲学认识主体分,可分为客观信息、主观信息;

按信息产生的时间分,可分为昔时信息、现实信息、未来信息;

按信息的应用领域分,可分为工业信息、农业信息、军事信息、政治信息、科技信息、文化信息、经济信息等;

按信息的逻辑意义分,可分为真实信息、虚假信息、不定信息、伪装信息;

按信息的产生效果分,可分为有用信息、无用信息、干扰信息,等等。

四、信息概念新论

从以上概述中可以看出,其实迄今为止,对信息的概念界定还是一个比较笼统、缺

乏统一的标准阐述,不论是香农、维纳,还是当今各学科对信息概念的界定,对信息的理解还是从各自专业角度去描述的。信息论、控制论、系统论和耗散结构理论是关于“信息科学”的最基础的四门学科,他们的“信息”定义也是有局限的,尽管在他们的学科领域是有效的,但都不是一个一般性的定义。近些年来,对信息概念的新研究、新探讨主要体现在以下几个方面。

1. 信息与存在的关系

钱学森曾经说过“信息就是信息,既不是物质也不是能量”。信息是一种存在,是事物之间的联系,这种联系可以是直接的,也可以使间接的。

2. 信息与语言、符号、知识的关系

语言的本体层面是本体化信息的一个环节,借助于存在的显现,语言成为作为存在者的语言自身,而信息则显示出双重性:它既通过存在的显现而显现自身,也经由语言的显现而在语言的境域中显现自身。符号可以视作一种语言,而知识是以为对象的信息劳动的价值及其积累,知识是在信息的循环过程中产生的。

第二节 文献信息基本知识概述及新论

一、文献的概述

“文献”一词在我国有着悠久的历史,最早见于《论语·八佾》篇中:“夏礼,吾能言之,杞不足之矣;殷礼,吾能言之,宋不足也。文献不足故也,足,则吾能之矣。”宋代朱熹注解为:“杞,夏之后;宋,殷之后,证也。文,典籍也;献,贤也。言二代之礼我能言之,而二国不足取之以以为证,以其文献不足故也。文献若足,则我能取之以证吾言矣。”其大意是指历史典籍与贤者。

最早以“文献”一词作为书名的是元代马端临所著的《文献通考》。他在该书的序言中,对“文”和“献”作了这样的解释:“凡叙事,则本之经史而参之以历代会要,以及百家传记之书,信而又者从之,乖异传疑者不录,所谓‘文’也。凡论事,则取当时臣僚之奏疏,次及近代诸儒之评论,以至名流之燕谈,裨官之记录,凡一话一言,可以订典故之得失,证史传之是非者,则采而录之,所谓‘献’也。”他认为,录之于书本的文字可称之为“文”,而传之于臣僚、诸儒、名流、裨官之口的言论则称之为“献”,从而继承和发展前人对“文献”的见解,缩小“文”与“献”的差别。

到了近、现代,文献的概念发生很大的变化。《辞海》(1979年版)认为只有图书文物资料才可称为文献,而贤者及其言论则不在文献的范畴之内,此时人们习惯于将“文献”解释为“专指具有历史价值的图书文物资料”。

近几十年来,随着科学技术的发展,记录知识的手段和载体也发生深刻的革命。从记录知识的手段来说,不仅有传统的刻、写和印刷,而且还有录音、录像和其他复制方式;从载体来看,不仅有传统的纸张,还有胶卷、胶带、录音带和计算机磁盘等。因此,“文献”一词也被赋予新的含义。《辞海》(1989年版)也对1979年版的定义作了补充:“今

又为记录有知识的一切载体的统称。即用文字、图像、符号、声频、视频等手段将人类的知识记录在各种载体(如纸张、胶片、磁带)上。”以此形成对文献的较为科学全面的解释。根据中国国家标准《文献著录总则》(GB 3792.1—83),文献是记录有知识的一切载体。国际标准《文献情报术语国际标准(草案)》(ISO/DIS 5127)则将文献定义为:“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,或存储在磁盘上。这种附着在各种载体上的记录统称为文献。”

从上述对文献的定义中可以看出,文献具有三个基本要素:一是要有一定的知识信息;二是要有负载知识信息的物质载体;三是要有记录知识信息的符号和技术。文献记录知识信息,而这些知识信息又依附于载体而存在。记录文献的载体有甲骨、金石、简牍、帛、纸、胶片、磁带、磁盘、光碟和在线网络等。

二、文献的属性

文献的属性主要包括以下四个方面。

(1) 知识信息性

这是文献的本质属性。任何文献都记录或传递一定的信息与知识。离开知识信息,文献便不复存在。传递信息、记录知识是文献的基本功能,人类的知识财富正是依靠文献才得以保存和传播的。

(2) 客观物质性

文献所表达的知识信息内容必须借助一定的信息符号、依附于一定的物质载体,才能长时期保存、积累和传递。

(3) 人工记录性

文献所蕴涵的知识信息是人们用各种方式记录在各种载体上的,而不是纯天然地荷载于物质实体上的。

(4) 动态发展性

文献并非处于静止状态,而是按新陈代谢的规律运动着的。随着人类记录水平的提高,信息交流的频繁,文献的数量日益庞大,形式日益多样;与此同时,文献的更新速度也在加快,生命周期日益缩短,形成有规律的运动。

三、文献信息

“文献信息”即文献中记录的知识信息,是指用文字、图像、符号、声频、视频等手段记录在一定的物质载体上的知识信息。它与“文献”并无本质上的区别。文献必须包含有信息,信息必须依附于一定的载体。文献信息是固化于载体上的知识信息,通过文献存储和传递,只有经过人们的利用,才会转化为动态的信息,发挥它的潜在的功能。人类社会利用文献或文献交流,实质上是利用和交流文献中的信息和知识。文献信息一般分为社会科学文献信息和自然科学文献信息。

1. 文献信息的功能

文献信息是科学知识存在和表现的形式,又是传播科学知识的基本手段,其功能主要有以下几个方面。

(1)加工功能

文献信息的生产和交流过程,并不是对各种自然信息、社会信息、生物信息等的简单再现和复制,而是需要不断地对信息选择、提炼、整序、转换等,这样才能更好地实现其传播、利用的价值。如一次文献信息、二次文献信息和三次文献信息。

(2)存储功能

文献是存储知识信息的最基本形式,是学习和研究工作的物质基础。虽然人们用以记录和存储知识的手段有大脑存储、实物存储和文献存储等,但是只有文献存储可用其具有人类共同的符号系统、多样化的记录技术和载体形式存储信息,扩大信息存储的范围和容量,延长信息存储的时间,以提高信息的使用率,实现知识信息的共享。

(3)传递功能

知识信息只有在传递的过程中才能发挥最好的功效,文献信息是传递知识信息的主要渠道。文献信息不仅可以在广泛的地理范围内横向传递知识信息,促进不同领域、不同地区之间人们的思想和学术交流,而且可以在较长的范围内纵向传递知识信息,将人类社会的精神财富代代相传。

(4)认识功能

人类认识世界、获得知识,一是靠亲身实践,获取直接经验,二是通过读书学习,获得间接知识。文献信息的产生和发展,使人们有可能不通过直接经验,而从阅览古今中外的各种文献中获取知识。这种间接获取知识的方法,可为人们从实践中获取知识奠定基础,由此拓宽视野,获取新知识,提高认识能力,加快对客观世界认识和改造的进程。特别是现代科学技术的进步,使作为知识载体的文献信息形态迅速发展,并日益多样化,这不仅扩大文献信息的外延,而且进一步提高文献信息的认识功能。

(5)参考功能

文献信息包含人们在各种活动中的思维进展、经验教训、研究成果和事实数据等各方面的有用信息。因此,它在人类各种实践活动中具有广泛的参考功能。如科研人员在科学研究的过程中,若以有关学术专著、论文、研究报告、文摘索引、综述评论、调查资料等文献信息为参考,就可避免不必要的重复研究,节省时间、精力、财力和物力,加快科研进程,提高科研质量。

(6)评价功能

文献信息是科学研究的重要保障,也是衡量、评价某人、某机构、某地区、某国家在某学科、某领域科研发展水平的重要指标。某人、某机构、某地区、某国家在某学科、某领域科研发展水平和地位,可以从其已出版或发表的文献信息中得到查证、确认和评价。

2. 文献信息的发展趋势

随着现在科技技术特别是网络技术的发展,文献信息在内容、形式、品种、数量等方

面都在发生着不断地变化,其发展趋势主要表现在如下几个方面。

(1) 文献信息的出版周期缩短,数量急剧增多

在信息社会,科学的发展、知识门类的扩充以及人们对知识和信息需求量的增加,导致各种记录人类科学知识的文献信息载体出版周期的缩短和出版数量的剧增。据不完全统计,20世纪70年代以来,全世界每年仅出版图书就达50万种,大约每一分钟就出版一种图书;到90年代,全世界每年出版图书已达到80多万种。随着网络技术的发展和人们网络阅读习惯的养成,网络出版物的出版速度更是惊人。

(2) 文献信息的内容交叉重复,分布异常分散

随着科学研究的不断深入,以及新兴学科的不断产生和发展,一方面各个学科之间相互渗透、相互融合,联系越来越紧密;另一方面,各个领域、各个学科不断分化、细化的现象也越来越多,这样就造成文献信息内容交叉复杂,专业文献信息分布又异常分散。同样的文献信息内容,往往会被各种不同形态的载体转载,也会被同一载体的不同方式转载。而专业文献信息也不全部是被本专业刊物登载,有相当部分的专业信息散布在其他专业的刊物中。这一趋势将使建立统一的文献信息资源体系成为必然,使文献信息资源的协调、收集、整理、规范等工作变得刻不容缓。

(3) 文献信息的语种范围不断扩大,译文数量剧增

目前,文献信息所使用的语种已越来越多,其中英文、德文、俄文、中文、法文、日文等所占比重较大,尤其是英文文献信息量约占全世界信息量的2/3左右。据统计,一个科学家在自己的研究领域至少有一半左右的文献信息量是用其所不懂的语言出版的。因此,为解决读者利用文献的困难,大量翻译外国文献信息已成为各国科学研究中的重要内容和手段。

(4) 文献信息类型复杂,载体形式多样化

人类社会在发展科学技术的同时,也在不断地发展记载科学知识的手段。现代社会的文献信息,除了拥有图书、报纸、期刊、会议论文、档案等印刷型文献信息外,还产生大量非印刷型、现代化的文献信息载体,如磁带、光盘、计算机网络,等等。而且,随着文献信息品种和数量的激增,文献信息的存储和检索必将日趋先进。

四、文献信息源

文献信息源是指人们获取文献信息资料的来源。文献信息源的形式有两种,即非文献形式和文献形式。非文献形式的文献信息源,通常是指没有印刷载体的文献信息源,如口头传播信息和实物传播信息,它们虽然在现代信息交流系统中占有重要的地位,也是人们获取知识信息的不容忽视的来源,但由于不能按原样存储和保管,所以不能把它们列为文献信息检索的对象。文献形式的文献信息源,是指通过有载体的文献资料形式而获得的知识、信息的文献来源,是图书情报界研究的对象,如常见的图书、期刊,等等。

按照不同的分类标准和文献信息存在的形态,可以把文献信息源划分为不同的类