

精·品·课·程·立·体·化·教·材·系·列



信息组织学 教程

周 宁/主编



 科学出版社
www.sciencep.com

国家精品课程 教育部规划教材



信息组织学

教程

第二版



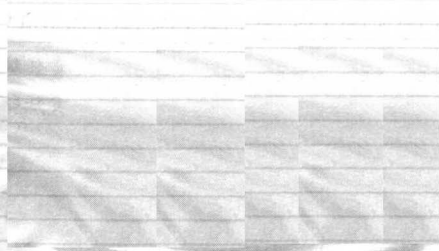
清华大学出版社
Tsinghua University Press

精·品·课·程·立·体·化·教·材·系·列

信息组织学

教程

周 宁 / 主编



科学出版社
北 京

内 容 简 介

本书系统地论述了信息资源组织的理论方法和技术,主要讨论了网络环境下数字化信息资源的描述、加工、存储、传递与利用的基本理论方法与实现技术。全书共9章,分别论述了信息组织学的基本原理与方法,信息资源管理的分类描述语言,信息资源组织主题描述语言,SGML、HTML、XML语言,元数据方案,数据库方法与搜索引擎等,以及信息组织在数字图书馆、电子商务系统、电子政务系统、数字博物馆与艺术画廊中的应用和发展趋势。

本书配备多媒体教学课件,适合作为信息管理类专业本科生教材,也可作为相关专业本科生和研究生的教学用书。

图书在版编目(CIP)数据

信息组织学教程/周宁主编. —北京:科学出版社,2007
(精品课程立体化教材系列)

ISBN 978-7-03-019284-4

I. 信… II. 周… III. 信息管理-高等学校-教材 IV. G203

中国版本图书馆CIP数据核字(2007)第097237号

责任编辑:林建 于宏丽/责任校对:陈玉凤

责任印制:张克忠/封面设计:耕者设计工作室

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

陈海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2007年7月第一版 开本:B5(720×1000)

2007年7月第一次印刷 印张:22 3/4

印数:1-3 500 字数:436 000

定价:30.00元

(如有印装质量问题,我社负责调换〈新欣〉)

前 言

科学发展、时代进步催生了新兴学科的诞生与发展。在人类从工业社会向信息社会迈进的过程中,信息科学中的一个重要分支学科——信息组织学得到迅速发展。

信息组织学是信息管理类专业的核心课程之一。它系统地讲述信息资源描述与组织的理论方法、基础技术与相关标准。随着 Internet 的迅速发展及普及,数字化信息资源已成为信息资源的主体。在信息生产与消费的过程中,人们对信息组织的理论、方法、技术等方面都有创新,并开辟了许多全新的应用领域。因此,本教材着眼于知识经济和信息社会发展的需要,重点讨论了以计算机为基础的数字化信息资源的描述与组织;吸收国际上该领域最新成果的内容,全面体现 21 世纪该领域的研究前沿的学术进展和主要成果,力争在体系结构、理论方法上有所突破。

本书由编者多年教案整理而成,是多年教学改革与创新的总结。同时,本书参考了多本国内外有代表性的教材,吸收了国内外的一些教学与科研新成果。

全书共 9 章,其主要内容如下:

第 1 章为导论,讲述了信息组织学的诞生与发展及其理论方法、技术基础和主要内容,并且讲述了信息组织在信息产业中的地位与作用。

第 2 章系统论述了信息资源管理的分类语言。它总结了国内外的主要成果和常用方法,并以网络搜索引擎为例,具体讨论了分类描述语言在信息组织中的广泛应用。

第 3 章讨论了信息资源组织的主题描述语言。从主题描述语言的诞生、发展

和广泛应用出发,具体介绍了国内外一些应用成果范例,进一步探讨了主题语言的构成与体系结构,列举了在网络环境下信息组织与资源共享的一些实例。

第4章系统论述了信息资源组织的语言工具。对 SGML、HTML 和 XML 的原理、相关标准与具体应用开展讨论,并就网上信息的组织与发布、多媒体信息的组织方法等从理论与实践的结合上进行研究。

第5章介绍了信息资源描述与组织的元数据方案。重点论述都柏林核心元数据研究的系列成果、定义、描述方法和具体应用。并对 RDF、PICS、CDF、MCF 等元数据与 RDF/XML 框架的应用作了进一步的研究。最后,讨论了(描述信息资源的)元数据的级别。

第6章论述了复杂元数据格式与信息资源组织的数据库方法。对 MARC21 (2709)、MARC21 (XML) 的理论方法和应用实例做了系统论述。并就相关的国际标准 (ISO 2709) 和国家标准 (GB2901) 格式与应用作了说明。信息组织的数据库方法经历了多种体系结构的发展和应用。对于新近出现的研究成果——半结构化数据库系统、多媒体数据库,也作了专题讨论。

第7章讲述了索引数据库与搜索引擎。重点对网络信息资源组织与利用的工具——搜索引擎作了系统论述。从搜索引擎的体系结构、建立与使用方法入手,结合国内外有代表性的搜索引擎作了具体研究,并对元搜索引擎进行了探讨。

第8章讨论了信息组织的广泛应用。针对数字图书馆、电子商务系统和电子政务中的信息组织、数字博物馆与艺术画廊等专题进行了探讨,从理论与实践的结合上进行了研究。

第9章论述了信息组织学的发展趋势。主要论述了信息构建、信息可视化、知识组织,讨论了信息组织与信息构建、信息可视化、知识组织三者的关系,并对上述三者的主要内容与具体应用分专题进行了探讨。

本教材中列举了大量的实例,力求理论与实践紧密结合;还列出了相关的实验课题供学员上机实践时参考。

本书按教学大纲要求,由集体编写而成。多年来,参加精品课程建设并撰写书稿的有周宁、文燕平、严亚兰、刘玮、杨峰、黄晓梅、张芳芳、余肖生、张会平;由周宁负责全书的规划设计与统稿。

在本书的编写过程中,作者得到了国内一些同行专家的关心、支持和帮助,参考了国内外许多学者的研究成果和论著,在此一并表示感谢。由于信息组织学课程涉及的知识面广、内容深、技术新,加之发展快,虽然我们付出了不懈努力,但疏漏之处在所难免,欢迎读者批评指正。

编者

2007年3月于武昌珞珈山

目 录

前言

第1章

导论	1
1.1 信息与信息组织	1
1.2 信息组织学理论方法	7
1.3 信息组织的技术基础	13
1.4 信息组织学研究的主要内容	20
1.5 信息组织在信息产业中的地位与作用	28
习题一	29

第2章

信息资源管理的分类语言	30
2.1 信息资源分类语言的沿革	30
2.2 分类语言的原理与体系结构	45
2.3 数字化信息资源管理中的分类描述语言	51
2.4 网站分类实例	60
习题二	66

第3章

信息资源组织的主题描述语言	67
3.1 主题描述语言的沿革	67

3.2 国内、外主题语言应用实例.....	84
3.3 网络信息资源组织中的主题法.....	94
3.4 信息组织与资源共享的实例.....	97
习题三.....	108
实验一 联机目录组织与信息资源共享.....	108

第4章

信息资源组织的语言工具	110
4.1 概述	110
4.2 SGML	111
4.3 HTML	115
4.4 Web 信息组织与发布	157
4.5 XML	160
习题四.....	192
实验二 网络信息资源组织.....	193

第5章

信息资源描述与组织的元数据方案	194
5.1 元数据的沿革	194
5.2 都柏林核心元数据	195
5.3 资源描述框架 (RDF)	216
5.4 其他元数据	219
5.5 元数据的级别	221
习题五.....	222
实验三 元数据描述.....	222

第6章

复杂元数据格式与信息资源组织的数据库方法	224
6.1 MARC21 (2709/XML)	224
6.2 ISO 2709 与 GB 2901	234
6.3 信息组织的数据库方法	237
6.4 半结构化数据库系统	245
6.5 多媒体数据库	247
习题六.....	249

第7章	
	索引数据库与搜索引擎 250
7.1	索引数据库与索引机制
7.2	网络搜索引擎的诞生与发展
7.3	搜索引擎的系统结构
7.4	搜索引擎的使用方法
7.5	搜索引擎的实例
7.6	元搜索引擎
	习题七.....
第8章	
	信息组织的广泛应用 274
8.1	数字图书馆的信息组织
8.2	电子商务系统的信息组织
8.3	电子政务系统的信息组织
8.4	数字博物馆
	习题八.....
第9章	
	信息组织学的发展趋势 321
9.1	信息构建
9.2	信息可视化
9.3	知识组织
	习题九.....
参考文献	 348
索引	 350

第1章

导 论

人类正从工业社会向信息社会 (information society) 迈进。在社会信息化、信息社会化的今天, 信息、物资、能源已成为现代社会的三大支柱。信息不仅是国家的战略资源、人类的宝贵财富, 而且信息的生产与消费水平已成为衡量一个国家综合国力和发展水平的重要尺度。在这里, 信息生产的核心是信息资源对象的描述与组织。

■ 1.1 信息与信息组织

1.1.1 信息与信息社会

信息 (information) 是音信、消息, 是物质存在的一种方式、形态或运动状态, 是标记在通信系统中的信号, 是构成知识的原材料。一般来说, 人们对信息有三种用法: 其一是作为事物的信息 (information-as-thing), 其二是作为知识的信息 (information-as-knowledge), 其三是作为过程的信息 (information-as-process)。信息泛指通过指令、数据、符号等发出的对接受者来说在此之前还没有获知的消息。人们获取信息、存储和传递信息、利用信息都需要信息技术 (information technology, IT)。它包括计算机技术、通信技术、传感技术和控制技术等相关的基本技术。目前, 信息技术高度发达并得到普遍应用, 形成了庞大的信息产业 (information industry)。

信息社会是高度信息化、知识化的社会。它的主要特点：在社会生活中广泛应用信息技术、信息获取与传递十分方便、信息生产发达、信息消费旺盛、促进了社会经济快速发展、人民生活质量迅速提高。

信息社会的形成经历了长期的发展过程，国内外专家、学者和有关的国际组织对信息资源开发、信息社会的形成和知识经济的发展作了大量研究，已取得不少成果。下面引述部分论述将有助于对其发展历程有一个系统了解。

1962年，美国经济学家弗里茨·马克卢普，根据美国从第二次世界大战以来至20世纪50年代末的社会发展和产业结构变化背景，提出了“知识产业”(knowledge industry)的概念。概念外延包括教育、R&D(研究开发)、传播业、信息设备及信息服务。

1973年，哈佛大学社会学家丹尼尔·贝尔出版《后工业社会的来临》一书，影响广泛，引起热烈的争论。早在1959年夏，在奥地利萨尔茨堡学术讨论会上，他已首次使用“后工业社会”一词；随后，逐步形成“后工业社会”的概念。他定义说，“前工业社会依靠原始的劳动力并从自然界提取初级资源”；“工业社会是围绕生产和机器这个轴心并为了制造商品而组织起来的”；“后工业社会是围绕着知识组织起来的，其目的在于进行社会管理和指导革新与变革；这反过来又产生新的社会关系和新的结构”。由于“后工业社会”的提法含糊，1979年，贝尔承认“信息社会”的概念较“后工业社会”更为确切。

1980年，未来学家阿尔温·托夫勒发表代表作《第三次浪潮》。他说：“我相信我们已处于一个新的综合时代的边缘。”他提出“超工业社会”的概念。关于“三次浪潮”的划分，渊源于贝尔。在其新著《权力转移》里，托夫勒预言：随着西方社会进入信息时代，社会的主宰力量将由金钱转向知识。同年，法国名记者和作家让·雅克·塞尔旺·施赖贝尔也发表《世界面临挑战》一书，他认为，信息是当今世界最重要而又取之不尽的资源，而自然资源与能源在地球上却日趋枯竭。书中他直接运用“信息社会”概念取代“后工业社会”和其他提法。

1982年，未来学家约翰·奈斯比特发表《大趋势》一书，从10个方面论述了美国社会发展趋势。他认为，“知识是我们经济社会的驱动力”，“信息经济社会是真实的存在，是创造、生产和分配信息的经济社会”。他概括了信息社会的四个特征，其中包括：起决定作用的生产要素不是资本，而是信息知识；价值的增长不再通过劳动，而是通过知识。

1985年，美国政府授权Calgary大学成立“知识科学研究所”(KSI)。把知识作为体系加以全面考察，研究知识对社会和经济等各方面的作用过程与转化过程。虽然尚未提出“知识经济”概念，但实际上已对知识经济的几乎所有方面做了富有成果的研究。

20 世纪 90 年代初,美国阿斯彭研究所(The Aspen Institute)等单位联合组建信息探索研究所(The Institute for Information Studies),在它出版的《1993~1994 年鉴》中,以“知识经济:21 世纪信息时代的本质”为总标题,发表了 6 篇论文,从 6 个方面审视了“明天信息社会”的特征和本质。在第一篇论文《技术在信息时代的地位:把信号转为行动》中明确地提出:“信息和知识正在取代资本和能源而成为能创造财富的主要资产,正如资本和能源在 200 年前取代土地和劳动力一样。而且,本世纪技术的发展,使劳动由体力变为智力。产生这种现象的原因,是由于世界经济已变成信息密集型的经济,信息和信息技术具有独特的经济属性。”

1994 年,C. 温斯洛和 W. 布拉马共同出版了《未来工作:在知识经济中把知识投入生产》一书,书中明确点明“知识经济”的概念,并对概念的内涵和外延作了较为完整的论述。作者认为,“管理智力”是获取和利用高价值信息的关键。他们阐述了知识形态下企业在市场取胜的基本条件和要求,并提出了“知识工人”(knowledge worker)的概念。

总部设在巴黎,以发达国家为主要成员国的“经济合作与发展组织”(OECD),在 1996 年 10 月 8 日发布了一份《科学、技术和产业展望》报告,在国际组织文件中,首次正式使用了“知识经济”(knowledge-based economy)这个新概念。OECD 将该报告的有关部分以“以知识为基础的经济”为题予以独立发表,文中对知识经济的内涵进行了界定:知识经济是建立在知识和信息的生产、分配和使用之上的经济。报告总结了 20 世纪 90 年代以来 OECD 国家经济发展轨迹与趋势,提出知识经济的主要特征及其作用、地位,用统计数字具体说明了知识经济体系中的重大要素,包括科学与技术的研究开发(R&D)、信息和通信技术、服务行业的就业人数与构成变化,以及劳动力的技能素质等因素对经济增长的影响。报告最后总结说:“事情已使人们越来越清楚:知识是支撑 OECD 国家经济增长的最重要因素。”

现代社会已进入知识经济时代,知识经济是和农业经济、工业经济相对应的一个概念。与以往的经济形态最大的不同之处在于,知识经济的繁荣不是直接取决于资源、资本、硬件技术的数量、规模和增量,而是直接依赖于知识或有效信息的积累和利用。知识经济的一个最直观、最基本的特征是知识作为生产要素的地位空前提高,知识不再是资本的附庸,它作为资源的最重要的形式不仅本身可以用来消费,而且它还能有效克服传统资源的瓶颈制约,并且能极大提高对传统资源的使用效率。

一般认为,信息是活动中的知识,同时,知识是经过整理和提炼的信息。知识是人们通过实践对客观事物及其运动过程和规律的认识,是信息的一部分,即被人们理解和认识并经过头脑重新组织和系列化的那部分信息,是经验、技能的

总结。在知识经济时代,社会与经济发展比以往任何时候都更加依赖于知识和信息。信息是构成知识的原材料,信息资源开发、信息生产与消费水平决定了知识经济的发展水平。

1.1.2 信息组织的诞生与发展

信息组织是对信息资源对象进行收集、加工、整合、存储,使之有序化、系统化的过程。组织信息的目的是为了检索信息、利用信息。图书馆、情报所、文献中心、档案馆、专利局等是人们长期进行组织信息和提供信息的场所,数据库和 Internet 是新的信息组织空间。

在信息生产与消费的过程中,信息组织是核心。作为社会活动的重要组成部分,信息组织是与人类社会同时产生并同步发展的。它经历了漫长的发展道路。在研究和应用实践中,可以用不同的方法来划分信息组织的发展阶段,按职能划分和按组织对象划分是常用的方法。

1.1.2.1 按职能划分信息组织发展阶段

信息组织的发展主要是与信息整体发展和信息组织理论方法发展及信息技术发展紧密联系的。因此,在发展阶段上,大致与信息管理发展阶段保持基本的一致性。由于在原始信息管理时期几乎谈不上任何形式的信息组织活动,信息的传递通常是不正规的、多自发性的。因此,在这里仅把信息组织的发展区分为三个阶段,主要与古代信息管理时期、近代信息管理时期和现代信息管理时期相对应。必须明确指出的是,信息组织发展阶段的演化并不是一个阶段对上一个阶段的全面否定,而是一种扬弃、一种理论和方法的完善和扩展。即使是最早的清册职能(inventory function)也并未在现代信息管理时期被彻底废止,而是融入了一种更为综合的组织职能之中。

1. 清册职能时期

在古代信息管理时期,信息管理的着眼点不在于“用”而在于“管”,信息管理活动并不要求完备、科学的信息组织给予支持。这一时期的信息组织活动主要体现为一种清册职能,它主要通过对信息的记录和登载,如同“信息账房先生”,告诉人们有什么信息而已。当然,这时的信息主要是指图书、档案等文献信息。偶尔人们也使用“catch word”等方式进行最简单的主题表示,分类法的使用本身也多用于信息编排,中国古代目录学家们的提要或提要式文字也并未改变总体的清册职能。显而易见,这种财产登记式的信息组织活动非常简单而不规范,手工方式和信息交流匮乏也决定了发展的不成熟性。这一时期的图书馆目录虽然还远远不能与今天的图书馆目录相比,却已成为这一时期的代表成果,围绕

图书馆目录展开的研究已初步显露了现代信息组织思想和方法的萌芽。

2. 查检职能时期

一切信息机构存在的目的都不可能是单纯的收藏和保管,利用才是最终目的。要促进信息的利用就必须有完备、科学的信息组织活动作基础。图书馆等文献信息机构的发展,特别是管理思想、技术和方法的发展,也为信息组织的发展创造了有利条件。因此,信息组织活动不仅告诉人们有什么信息,而且告诉人们怎样才能找到信息,构成这一时期的信息组织活动的主要职能就是查检职能(finding list function),“信息账房先生”摇身一变成了“信息药房先生”。但仍然可以发现,这里的信息还没有冲破文献信息的限制,现代意义上的信息还缺乏有效的、有意识的控制活动。但是,信息组织的理论、技术和方法已经不可同日而语了。信息组织已经不再是一种记簿式管理,尽管对信息的概念单元分析仍然存在不足,但对信息的外在特征和内容特征的描述和表示都有不同程度的重视。信息组织不仅能够为用户提供更多的检索点(access points),而且同时,索引和文摘的出现使二次信息体系中书目形式一统天下的局面被彻底打破。信息组织的纯手工操作方式也不复存在,机械化、半机械化手段开始导入信息组织中,出现了穿孔卡片系统和缩微胶片系统等新的信息组织技术手段。不仅信息描述内容丰富、信息表示方法多样化,而且信息组织从载体单元到内涵单元的重点倾斜具有里程碑意义。无序信息向有序信息的转换使信息的利用有更多的方法和途径,在近代信息管理时期,信息组织确立了“用户中心论”,并在检索点的形成和排列、整理成果的推广和应用、传统技术的更新和演化等方面得以充分体现。在这一时期,信息组织的发展可以用主题揭示语言对分类揭示语言的冲击、文摘和索引型有序化成果对书目型有序化成果的冲击、机械化和半机械技术手段对手工技术手段的冲击,给予充分地概括,这些变化特征为信息组织由查检职能为主导向完整意义的组织职能发展奠定了坚实的基础。

3. 组织职能时期

以计算机技术、通信技术、网络技术为代表的现代信息技术的导入,把信息管理推向一个新的发展水平,也把信息组织引入一个新境界,使信息组织过去连想也未曾想到的新概念和新设想变成了现实。信息组织从最完整的意义上实现了整理信息、科学分流、促进选择、保证利用的组织职能(organizing function)。这种组织职能已经不满足于告诉人们有什么信息和怎样找到信息,它不仅可以告诉人们这里有什么信息,而且还可以告诉人们其他地方有什么信息;人们不仅可以找到最需要的信息,而且还可以知道什么信息适合于自己;人们不仅可以利用一种途径获得信息,而且可以多途径、多角度地查检信息。在这里,信息的概念被高度扩充,不仅有文献信息,而且有非文献信息;不仅有文字信息,而且有声音、图形、动画或视频信息;不仅有知识信息,而且有满足非学习研究型需求的

其他信息；不仅有提供线索的二次信息，而且有直接满足要求的一次信息和三次信息；不仅有被动接受的信息，而且有人们参与生成和选择的信息……所有这一切，反映了现代信息组织的超前性和现实性、精细性和普及性、协作性和专业性、标准化和多样化、体系化和特色化、市场性和调控性、自动化和智能化的有机结合。信息组织融信息搜集、信息分析、信息描述、信息揭示、信息存储于一体，现代信息管理时期中信息组织活动的社会化、信息组织产品的商品化、信息组织需求的市场化成为这一时期最明显的标志。在未来发展的道路上，信息组织的功能不会弱化，而且会进一步强化。可以预言，信息组织、信息传播、信息生产与消费这几者之间的联系会更加紧密，一体化趋向会更加明显，但其中的信息组织的功能将依然存在并得到充分体现。

1.1.2.2 按组织对象划分信息组织发展阶段

IT 的进步已成为现代信息组织发展的内在驱动力；同时，在社会信息需求的推动下，现代信息组织方法在不断地推陈出新。

1. 文献组织阶段

传统的文献组织是伴随着文献数量的激增而诞生的。近代，随着科学技术的不断发展，科技文献在出版种类、出版数量、出版速度等方面以飞跃式的速度向前推进；文献分布极其分散，文献语种迅速扩大；各种文献之间彼此交叉、重复日益严重；文献新陈代谢迅速加快。与此同时，传统学科界限不断被打破，学科越来越多、越来越细，这些都导致对于文献信息需求的不断深入和集中，使得文献组织工作显得尤为重要。传统意义上的文献组织，事实上是将文献由一次到三次的转化过程。在文献由一次信息向三次信息的流动与演变过程中，不断加入了文献著者、编辑者、信息加工者的创造性劳动，逐步使得文献所含知识得到鉴别、提炼和综合，从而使文献流不断增值。同时，也使得信息由分散到集中，由无组织到系统化。

这一阶段主要以图书情报领域的信息组织工作为主流。在传统的信息管理中，图书馆的工作核心是文献组织，即图书馆馆员根据文献的形式特征和内容特征，通过编制目录、索引、文摘等方法，对文献内容进行概念逻辑划分，将成千上万的文献按其标识排成一个有序的系统，向人们提供手工检索或计算机检索的过程。实质上，文献组织仅是对文献的形式上的加工，而没有对文献的内容进行组织，也就是说，仅对信息的载体进行组织，而没有组织信息本身，属于一种初级的信息组织。

2. 信息组织阶段

随着信息技术的发展，信息资源的载体、处理手段等信息的存在环境和条件正在发生前所未有的变化，并有加速的趋势。这种趋势促使信息组织的对象、形

式和内容都经历着根本性的变异,以纸介质为主的信息记录和组织方式的重要性正在下降,信息组织的中心点正转向网络信息资源和电子介质,特别是转向自动化、功能全、速度快的信息资源组织方法。人类信息活动正从以纸介质信息为中心转向以电子媒介为中心,这种走向推动着信息组织由传统方式向现代方式转变。

随着互联网的出现及其迅猛的发展态势,网络信息在社会信息量中所占的比例呈日渐上升的趋势。针对网络信息数量庞大、类型繁多、难以规范等特点,专业研究人员开始探讨传统的分类法、主题法以及分类主题一体化方法在网络环境下的转换,同时研制出了适用于网络信息组织的搜索引擎。此外,还开发出了一系列适合于网络信息资源描述与组织的工具——元数据。

3. 知识组织阶段

知识组织(knowledge organization)一词,最早是于1929年由英国著名图书馆学家布利斯(H. E. Bliss)提出来的。布利斯于1929年就出版了《知识组织和科学系统》、《图书馆的知识组织》两部著作,阐述了从文献分类角度组织知识的思想。但对知识组织进行专门的深入研究却只是近十年的事情。知识组织是指为促进或实现主观知识客观化和客观知识主观化而对知识客体所进行的诸如整理、加工、引导、揭示、控制等一系列组织化过程及其方法。

信息是产生知识的原料,信息组织对信息的有序排列是知识组织的原料基础,知识组织的出现不是一蹴而就的,而是信息组织发展到一定阶段的必然产物。信息组织是将处于无序状态的特定信息,根据一定的原则和方法,使其成为有序状态的过程,以方便人们利用信息和有效地传递信息。其提供给用户的是片面的、零散的信息,用户要想加以利用,还必须进一步对其进行分析、选择、加工。而知识组织的目标是对知识存储进行整理和提供知识,它提供给用户直观的、可操作性强的系统化知识。知识组织不再局限于利用片面的信息来满足用户的需求,而是对用户的需求进行系统分析,向用户提供全面、完善的解决方案,用户可以直接学习这些知识而无需再加工。

对于人类社会的发展和人类文明的延续而言,信息组织不可或缺,知识组织才是最终的目标,后者以前者为基础,而前者以后者为发展目标。

1.2 信息组织学理论方法

如前所述,信息组织是对信息资源对象进行收集、加工、整合、存储,使之有序化、系统化的过程。而信息组织学是系统研究信息组织的理论、方法和技术的科学。

1.2.1 信息组织的理论基础

信息组织是由来已久的一种人类社会的实践活动,在其发展过程中,不断从相关学科的理论和方法中汲取营养,使自身逐渐得到充实和完善。

系统论、耗散结构理论和协同论是信息组织的理论基础。

(1) 系统论。系统科学的思想是 20 世纪 20 年代由奥地利学者贝塔朗非(Ludwig Von Bertalanffy)提出来的,他把系统定义为相互作用的诸要素的复合体,认为系统的定义可以确定为处于一定的相互关系之中,并与环境发生关系的各组成部分的总体。系统论认为,系统内部各要素不是一种简单的结合,而是相互关联地有机结合在一起。正是由于系统要素之间这种有机的相互关联关系,系统的整体功能才产生了质的飞跃,远远超出各单个要素的功能总和。这就是系统论的“整体大于部分之和”的原理。在信息组织中,如果将大量的、分散的、杂乱的信息组织成一个系统,建立起内在的关联性,那么,信息系统的整体功能将大于各个信息单元的功能的总和。也就是说,这将能充分发挥信息资源的价值与作用。基于这一原理,信息组织的目标是要建立信息系统。

(2) 耗散结构理论。20 世纪 60 年代,比利时学者普里高津(Ilya Prigogine)提出了一个著名论断:在开放系统中,系统不断与外界进行物质和能量的交换,熵趋于最小值,能量远离平衡,混乱度最小,从原来无序结构转变为一种时间、空间和功能上的有序结构。这就是适于一切开放系统的耗散结构理论。耗散结构是一种自组织结构。自组织的机理:系统内各组成部分或要素之间有着强烈的相互作用,由此产生多方面的相互影响、制约的相干状态,致使各级组织不断地与外界环境进行交换,从而使自身保持稳定,在动态中实现自身的有序结构和特定功能。这样使得系统在非平衡状态下,实现的是动态、有序的活结构,而不是孤立、趋于无序的死结构,也不是变化微小、不能产生新组织结构的近平衡结构。普里高津认为非平衡、开放系统是有序之源。这一原理给信息组织的启迪:所要建立的信息系统应该是一个开放系统。现实存在的包括信息组织和信息检索在内的信息系统正是这样一个开放系统,它与外界进行着信息的交换,既采集信息也输出信息。

(3) 协同论。协同论是斯图加特大学物理学教授海尔曼·哈肯(H. Haken)创立的,以其系统演化方程为中心内容。这一内容有:自然界事物是多因素体系,即复杂系统;系统的总状态取决于作用,至少非线性的驱动力和涨落力,即系统所产生的扰动是不可缺少的;注重系统内在要素的相互作用的同时,环境的影响也不能忽视,前者主要表现为涨落,后者主要表现为外控制参量;时空也是影响系统状态的要素,因为运动是主体的运动、历史的运动。由于系统内各要素