

经济信息概论

燕惠兰 主编

航空工业出版社

经济信息概论

主 编 燕惠兰
副主编 刘晓敏

航空工业出版社

1998

内 容 提 要

本书全面系统地论述了社会主义市场经济体制下经济信息和其他社会信息的运动规律,组织、交流与传播理论以及信息服务、信息市场、信息分析与应用、信息系统的开发与管理等方面的理论与方法。既可作为经济、贸易、工商管理、经贸信息和情报信息专业本科教材,也可供广大科技工作者、经济贸易企业家和情报信息研究等部门有关人员参考。

图书在版编目(CIP)数据

经济信息概论/燕惠兰主编. —北京:航空工业出版社. 1998. 1

ISBN 7-80134-165-1

I. 经… I. 燕… III. 经济信息学 N. F208

中国版本图书馆 CIP 数据核字(97)第 24655 号

航空工业出版社出版发行

(北京市安定门外小关东里 14 号 100029)

北京地质印刷厂印刷

全国各地新华书店经售

1998 年 1 月第 1 版

1998 年 1 月第 1 次印刷

开本:787×1092 1/16

印张:15.25

字数:378 千字

印数:1—2500

定价:16.50 元

前 言

《经济信息概论》是高等院校信息管理类专业的一门重要专业课程。在几年的教学实践中,我们深感国内缺乏这方面的系统、完整而且适用的教材。因此,在教学的同时,我们开始搜集积累资料,并尝试着将自己在教学过程中形成的若干思想进行归纳和总结。1995年航空工业总公司教育培训部将《经济信息概论》列为总公司“九五”规划教材,由此我们正式开始了本教材的编著工作。

信息是当今社会一个相当普及的术语,它已深入到社会的各个阶层、每个角落,没有人不知道它的魔力、它的价值。然而,它又是令我们长期以来感到困惑、并将继续困扰我们的名词。当我们试图拨开笼罩其上的团团迷雾,深入到内核认识和描述其真实面貌时,又不得不发出“画鬼容易画人难”的感叹。信息是无限的,对信息的认识也是无限的。本书的框架仅限于对经济信息和其他社会信息的运动规律,组织、交流与传播理论以及信息服务、信息市场等方面。至于自然信息的规律及其具有普遍意义的信息传输、处理、识别等方面的问题,则属于包括信息论、控制论、计算机科学与通信科学在内的广义信息科学的范畴。

在编著过程中,我们力求做到体现以下特点:一是及时反映信息工作的实践经验,并对其进行总结与归纳。最近十多年以来,信息产业作为一个新的产业在逐渐兴起,在国民经济中发挥着越来越重要的作用。信息工作、信息服务正在由事业型向产业型转变,信息市场开始形成。信息工作的实践要求相应的理论指导,许多研究人员对此进行了富有成果的探索,对这些成果,编著者尽量加以提炼和系统归纳。二是跟踪信息技术的发展。信息技术是信息工作的物质基础,代表了信息工作的水平与层次。信息学的发展也证明信息学理论体系是某一特定时期信息技术水平的反映。当前随着以计算机技术为核心的信息技术的飞速发展,传统意义上的信息、文献、信息交流、信息系统等概念已发生了深刻变革,电子信息等新型载体信息的日益普及,对信息学提出了许多新的研究领域。最后注意吸收信息学理论研究的最新成果并使之在各章内容中得到充分体现。

本书分为八章,各章的具体分工是:第一、五章由燕惠兰撰写,第二、七章由刘晓敏撰写,第四、六、八章和第三章第四节由郭秋萍撰写,第三章由刘焕成撰写。全书由燕惠兰、刘晓敏最后修改定稿。编著过程中,我们引用、参考并借鉴了国内外许多学者的学术思想与观点,全书承蒙王人骅教授审校,并提出许多宝贵意见,在此,一并向所有给予我们指导和帮助的专家学者及同行们表示诚挚的谢意。同时,我们也自知由于受编著时间、研究水平等因素的限制,本书中还有着缺陷和疏漏,恳请专家、读者批评指正。

编 者

1996年11月

目 录

第一章 信息学基础理论	(1)
第一节 信息学发展史.....	(1)
第二节 信息学研究领域.....	(3)
第三节 信息学基本原理及理论体系.....	(8)
第四节 信息的量与质	(11)
第五节 经济信息学	(16)
第二章 经济信息源及其传播	(21)
第一节 经济信息源概述	(21)
第二节 经济信息源的分布规律	(24)
第三节 经济信息传播	(34)
第四节 经济信息传播方式	(43)
第三章 经济信息的搜集与整理	(53)
第一节 经济信息的搜集原则和程序	(53)
第二节 经济信息搜集的渠道和方法	(57)
第三节 经济信息整理	(62)
第四节 经济信息的组织方法	(73)
第五节 经济信息存贮	(80)
第四章 经济信息检索	(88)
第一节 经济信息检索的原理与方法	(88)
第二节 经济信息检索的数学模型	(91)
第三节 经济信息的手工检索	(95)
第四节 计算机经济信息检索系统	(99)
第五节 经济信息检索系统的评价.....	(104)
第六节 经济信息检索新技术.....	(106)
第五章 经济信息分析与应用	(110)
第一节 经济信息分析与预测的原理与要素.....	(110)
第二节 经济信息分析的基本程序.....	(112)
第三节 经济信息分析方法.....	(115)
第四节 定性分析方法.....	(117)
第五节 探索型预测方法.....	(126)
第六节 时间序列分析法.....	(131)
第七节 季节预测.....	(135)
第六章 信息市场	(145)
第一节 信息市场概述.....	(145)

第二节	信息商品	(148)
第三节	信息市场体系和结构	(153)
第四节	信息市场的交易形式	(156)
第五节	信息市场的管理	(160)
第六节	信息用户	(163)
第七章	信息产业	(170)
第一节	信息产业概述	(170)
第二节	信息产业的部门分析	(182)
第三节	信息产业的测算方法	(187)
第四节	信息产业管理	(195)
第八章	经济信息系统的开发与管理	(201)
第一节	经济信息系统概述	(201)
第二节	经济信息系统的开发方法	(206)
第三节	经济信息系统的总体规划	(211)
第四节	经济信息系统的分析与设计	(213)
第五节	经济信息系统的实施、运行和维护	(219)
第六节	经济信息系统的评价	(221)
第七节	经济信息系统的优化与控制	(223)
第八节	经济信息网络	(226)
第九节	国家信息政策与法规	(231)
参考文献		(237)

第一章 信息学基础理论

第一节 信息学发展史

信息学的产生,是人们长期实践、总结、升华的结果。信息科学的发展也经历了实验性描述、理论综合、数理论证各个阶段而日臻完善。作为一门新兴的综合性的横断学科的信息学一经诞生,就以其强劲渗透力,迅速渗入到社会的各个角落。它的纽带作用和催化作用,有力地促进了各相关学科的发展,日益显示出它的深远影响和强大生命力。信息科学技术的广泛应用,不仅促使社会生产力产生巨大的飞跃,而且还带来了产业结构、生产管理等方面的重大变革。人类正进入以开发信息资源为主要特征的新时代。

一、信息的起源与人类信息活动实践

信息的最初起源可以追溯到宇宙之初。早在人类产生之前,宇宙中的射电源就不断地向宇宙空间发射电波,这种射电波就是射电源存在的信息。世界万物,向外界发射各种信息,藉以表示自身的存在和运动。日月星辰表明宇宙的存在;枝头发绿,报道春天的来临;各种虫鸣鸟叫是呼唤同伴的语言;蝴蝶蜜蜂的舞姿是它们行动的信号。自从有了大自然,就有了自然、生物信息,这些信息的产生和存在,是人类探索大自然,开发和利用自然资源的重要依据,也为人类社会的信息活动提供了宝贵的信息资源。

自从有了人类,也就有了人类的信息交换、信息利用的活动。人类对信息的识别、获取、处理和使用是在实践中不断认识和提高的,并经历了一个漫长的过程。早在 300 万年前,处于原始状态下的人类,只能靠自身的感觉器官接受自然信息,并用声音、手势、表情等来传递一些极其简单的信息。劳动使原始人类的交往日益密切,互相帮助和共同协作的场合逐渐增多,“这些正在形成中的人,已经到了彼此间有些什么非说不可的地步了”,于是语言便产生了。语言是人类传递信息最基本的工具。人类藉此可交流劳动中获取到的丰富多样的信息,并准确地表达喜怒哀乐等复杂的思想感情。语言的产生既促进了生产力的发展,同时也促进了人脑存贮与处理信息的功能,它是人类传递与利用信息的第一次飞跃。

随着生产的发展,引起了体力劳动和脑力劳动的分工,也促进了文字的产生。在 6000 年前,人类发明了文字,用以记录、保存和传递信息,使人类获得了比语言更为有力的记载和传递信息的工具。从此人类开始了有意识的信息活动。公元前 25 世纪,古埃及人把纸草(papyrus,一种植物,英文 paper 来源于此拉丁文)的茎刻成薄片,压平后用作书写的材料;若干片粘成长幅,卷在木杆上形成卷轴,这就是在古埃及发现的最早的书籍。后来,古希腊人在卷轴上系上标题(titulus,即英文的 title),并按字顺排列。公元前 1700 多年,在巴比伦国王汉谟拉比时期,人们把楔形文字写在粘土板上。公元前 1400 年的我国殷代甲骨文,已经是一种具有严密规律的古文字。文字的产生是人类传递、利用信息的第二次飞跃,它使信息的交流活动突破了时空限制,得以世代相传,从而把人类带进了文明时代。

早期的信息载体多为龟甲、兽骨、金石、简策、布帛等。公元 105 年,我国东汉时期的蔡伦发

明了造纸术,解决了记录文字的载体问题。1041~1049年间,我国宋代的毕昇发明了胶泥活字字板,1450年德国人古登堡(J. Gutenberg)综合了中国油墨技术和朝鲜的铜活字技术,发明了活字印刷术。从造纸到印刷技术,使文献成为科技交流的主体。

从文献工作角度出发,人类利用文字记载、传播信息的实践可以追溯到遥远的古代。公元前7世纪的亚述班尼佩尔(Assurbanipal)时期就有刻在粘土板上的目录。我国春秋末期的孔子在整编六经时,就创立了揭示文献提要的方法。希腊的伟大哲学家亚里士多德(公元前384~前322年)是目录学理论的创始人。他编写的第一部百科全书,是一本综合性分类学的著作。我国西汉刘向、刘歆父子经过20多年的工作,完成了《别录》和《七略》,这是我国最早的书目编撰,也是世界最古老的大型综合目录。以后,班固(公元32~92年)编撰的《汉书·艺文志》在分类、目录、提要等文献工作中,做了许多有意义的工作。

现代图书目录被认为始于1564年,威勒(G. Willer)出版了书本式目录。1654年,盖斯勒(K. Gessner)编辑了《世界目录大全》(Bibliothca Universalis)。最早的科学杂志是1665年在伦敦出版的《哲学汇刊》(Philosophical Transactions)和法国出版的《学者杂志》(Journal des Sarants)。

18世纪,文摘期刊在德国取得了进一步发展。1751年出版的《医药图书目录》,是最早的专业性文摘期刊。1765年又出版了《德国图书目录》。

目录学是在人们大量从事书目实践的基础上逐渐形成的。18世纪大量文摘期刊的出现,使目录学有了进一步发展。

近代社会,随着科学技术水平的发展,人们在研究与利用电磁波方面获得了重大突破,发现电磁波可以携带信息。于是先后发明了电报、电话并架设了广播、电视。从此,信息传递方式发生很大变化,人们不但可以超越空间在瞬间同时收听语言和音响信息,而且可以看到图像和文字信息,从而空前地提高了信息记录与传递的速度和准确性。第二次世界大战后,录像、录音又相继发展起来。60年代以来,现代科学技术的迅速发展,通信卫星,微电子技术和计算机的广泛应用,使信息传递手段迅速增加,70年代后,报纸传真、微波通讯、激光电视及光盘等多媒体技术的问世,使信息的传递速度加快,同时也打破了地区界限。遍布世界各地的信息网络的形成,开辟了信息收集、存贮、处理、传递的新时代,人类开始进入了信息化社会。这是人类历史上利用信息的第四次飞跃。

二、信息学的萌芽和形成

信息作为一个含义相当广泛的概念,已经渗透到自然界、人类社会和人类思维的各个方面。随着信息理论的不断发展和完善,展现在人们面前的信息知识越来越丰富,它所涉及的范围也越来越广泛。

追溯信息学的萌芽,应当源于“科学信息”这一术语的提出。在1945年6月公开发表的给美国政府的信件中,布什(V. Bush)总结了战时科学信息的形成和发展,并对信息管理与应用提出了新的技术设计。即在应用缩微技术、存贮技术、代码技术的基础上,设计了“Memex”装置,提出采用机械化缩微检索方法查找文献。科学信息的概念后来被许多学者所接受,并与科学交流相联系,发展成信息学的交流学派。

1950年,在一次国际数学会议上,莫尔斯(C. N. Mooers)提出了“信息检索”这一术语,给信息查询活动以科学的概念。

“信息检索”、“科学信息”这两个学术概念的出现推动和发展了信息学学科的形成和建立,被认为是信息学诞生的先声。

1948年,申农(C. Shannon)和维纳(N. Wiener)先后发表了《通信的数学原理》和《控制论:或动物和机器中控制和通信的科学》两篇著作,为信息学的建立奠定了理论基础。同年,在伦敦召开了由各国学者参加的第一次国际科学信息学术会议。1958年由美国文献学会(ADI)、美国科学基金会、美国科学院等联合召开了华盛顿国际科学信息会议。这两次重要的会议为信息学的建立打下了牢固的基础。

申农的信息论理论一经问世,便改变了科学界的想像,给许多学科带来了新的希望。于是,人们试图用信息论的概念和方法解决各自面临的各种难题,诸如组织化、语义学、生理学、心理学、人类学、经济学、统计学等。信息论的冲击波便迅速冲击各个学科,人们尽可能从广泛的意义上理解和解释信息论。50年代至60年代,人们对信息的实质进行了一些深入的研究,使申农信息论在概率信息范畴内日益成熟起来。60年代中后期到70年代,由于科学技术的交叉与渗透,信息论中的一些概念和基本理论已经通过通信领域,得以在其他学科领域中推广和应用,这就使仅局限于通信技术领域的申农信息论暴露出明显的缺陷。针对申农信息论未考虑信息的内容及接收者使用价值等不足之处,人们先后提出了无概率信息、定性信息、偶发信息、模糊信息、相对信息等一系列新的概念和方法。相应地,各种语义信息和语用信息的描述和度量方法也先后出现。至此,人们无论对于信息的本质及其度量方法,还是对信息运动变化的一般规律都有了一定程度的探索。于是广义信息论即信息科学便在70年代初应运而生了。

信息科学是研究信息现象及其运动规律和应用方法的综合性科学,是以“三论”为理论主体,以电子计算机等为主要工具的一门新兴的横断学科。信息科学涉及与信息有关的一切领域,如计算机科学、仿生学、人工智能等学科,但并不是包罗万象的科学之科学。信息学已经发展成一门对社会进步举足轻重的新学科。

第二节 信息学研究领域

信息学是一门新兴的学科,正处在形成、发展与完善之中。信息学与其他学科的根本区别在于它有着确定的、特有的研究对象——信息与信息过程。

一、信息的基本概念

目前人们对信息的定义认识还不一致。就目前资料,世界上至今还没有一个公认的信息定义,而只能从不同的角度出发,对信息的各种现象和活动进行不够完善的定性描述。

对“Information”的解释有几十种,直接与知识信息有关的定义有37种。早在1959~1971年间,联合国教科文组织(UNESCO)负责科技信息的机构PGI就对37种定义进行过研究。这些定义的释文由词和词组组成。这37种定义共使用自然语言词汇81个,如知识、应用、技术、描述等。词频统计的结果,使用频率最高的词依次是:“传递(Transmission)、语言、流(Flow)、处理、交流(Communication)。在37种定义中,兼容这5个高频词汇的定义只有3个。

80年代初,我国学术界对“信息”的定义进行过十分热烈的讨论,虽未产生一个被共同接受的定义,但为信息学的研究开拓了崭新的领域。

事实上,自50年代以来,“Information”一词使用十分广泛,几乎所有学科、专业,乃至日

常生活的各个方面都使用它。主要原因是,它涉及到科学思想的发展、技术发明与应用以及科学技术进步造成的社会变革等,因而相关的术语、学术观点、科学理论也在迅速发展。虽然目前对于“知识信息”的概念尚未取得统一的认识,但其中三点是共同被强调的:

1. 为了一定目的而系统地收集起来的正确的信息和事实所构成的知识,也就是说,经过传递而又为人们所吸收的知识构成信息。

2. 通过传递而进入人类社会交流体系的运动着的知识构成信息。即一方面信息可以通过一定的物质载体或其他交流形式进行传递,另一方面获取信息必须经过传递。

3. 信息不仅是传递中的知识,而且必须是有效用的知识。凡是人类社会的信息均有一定的效用,或与接收者原来知识状态吻合而被认可,或有相异,从而发生知识的重构,生成新的知识结构。

因此,在人类社会交流过程中,知识变为信息,信息转化为知识,它们总是处在相互作用之中。知识是静态的,以实践为基础,通过抽象思维,对客观事物的规律性加以概括和总结。知识信息是人类社会中客观存在着的部分社会信息,是以知识形态为主,包括数据、消息等非知识形态的社会信息。所以说,知识信息是知识的激活,知识是信息的条理化。

“Information”一词来源于拉丁语,表示传达的过程和内容。最初传达仅仅意味着告知,而被告知的只能是以语言为外壳的人的思想内容。也就是说,“Information”是人类社会特有的现象。随着科学技术的发展,“Information”的语用范围也随之扩大了。1948年,N. 维纳以《控制论:或动物和机器中控制和通信的科学》为标题,阐述控制论的基本原理。后来,维纳关于“Information”的论断被广泛引证,人们开始把“Information”与诸如空间、时间、运动和能量之类的普遍概念联系在一起。这样,Information(信息)不是一个概念,而是一组概念群(见图 1-1)。N. 维纳说:“信息就是信息,既非物质,也非能量。”⁽¹⁾美国学者萨拉塞维克说:“在研究信息时,我们面对着一个展示出其物理的,生物的和行为的多种特性的复杂现象。”⁽²⁾他把问题引向了物理学、生物学和行为科学的多种不同层次上。英国学者布鲁克斯(B. C. Brooks)将信息看成是知识结构的一个小小的组成部分,它可以改变知识的结构⁽³⁾。我国信息科学家钟义信教授说:“信息是事物存在方式或运动状态,以及这种方式(状态)直接或间接的表述。”

由以上学者所下的定义可以看出,信息处于一个多维与多层次的定义域里。

在我们看来,科学的做法,应从“纯客观”及“使用者”两个不同角度来分别考虑其定义。从“纯客观”角度来说,“信息就是事物运动的状态和方式”,也就是说,一切事物(自然界的、社会的、思维领域的)都在运动,这种运动状态和运动方式就是信息。这里,状态是事物运动中相对稳定的一面,方式是它变动的一面。在“纯客观”的定义之外,还必须把信息的使用者(即用户)考虑进来。从用户的角度看,信息就是关于事物运动方式(状态)的广义知识。当然,使用者从实用角度对信息的理解可能出现很大差异。例如,在

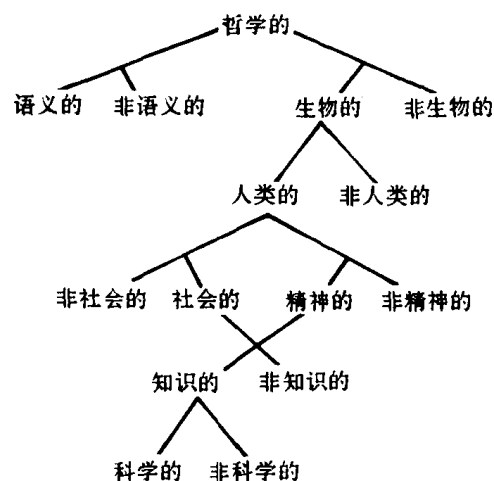


图 1-1 信息定义域

情报学界,信息一般被理解为知识情报;在遗传学中,信息往往被理解为 DNA 的结构形式、组织方式来定义;在管理学界,信息则被理解为管理活动的特征及发展情况的知识、情报、资料等的统称。

由此看来,信息不仅存在于自然界、社会,而且也存在于思维领域。哪里有事事物(物质的或精神的),哪里就有信息。信息是人与外部世界的中介,是人与外部世界联系的纽带。因此我们认为:广义的信息是物质和能量存在与运动的标记,它反映着物质和能量在空间和时间中的分布状况及其发生变化的程度,是各种系统普遍联系的中介。它不仅出现在人类社会中,而且普遍存在于宇宙之间,甚至存在于生物个体内部各器官之间,各细胞之间,各生化成分之间。

人类社会的信息是以语言、文字、图像、数据和状态为载体的。例如,体温计的读数告诉医生有关病人体温高低的信息;电子秤显示的数字是报告物体重量的信息;电报、电话、交谈、信件是人们相互传递生产、生活、学习、工作各方面的信息;书刊报纸则向人们传递各类知识和新闻信息。

事实上,任何人从事任何工作,都需要信息。领导者制定方针政策,作出决策,发布指示,实际上就是信息收集、处理和利用的过程。工业生产从厂房建设、设备安装、生产设计到制造工艺,都需要利用信息来执行管理、调度、计划和检查。农业生产要了解种子、化肥的供给情况,预测天气、水土的变化而实行科学种田。商业经营更离不开市场需求、消费水平、价格变动、仓储运输、顾客心理等信息。总之,在整个人类社会,到处都充满着信息和信息交流,人们的一切活动都离不开信息,人类生活在信息的海洋中。因此,我们认为:信息是一切事物、物质某种属性的反映,人们通过它可以了解事物的或物质的存在方式和运动状态,它最普遍的形态是信号、消息、数据、事实、知识、情报;它可以交换、传递和存贮,是一种能够创造价值的资源。

二、信息学有关概念体系

信息学以自然和社会中普遍存在的信息现象和信息过程为研究对象。我们在这里将信息学作为术语来探讨其概念体系。

术语学(Terminology)是在一个专门主题领域里,或所有主题领域的集合里,涉及概念的构成与命名的科学。术语(Term)则是表示专门知识领域的概念,反映概念相互关系的词或词组。因此,从术语学的观点出发,术语是研究表达概念所需的名称(即称谓),而不是术语的词或词组的名称所表示的意义。因此,研究术语要划分和确定概念并给予明确的定义,分清概念的范畴和关系,然后将概念按学科进行分类。本节从术语学的观点出发,对信息学的概念体系作必要的阐述。

(一)外来各种信息学名词辨析

1. Information Science(信息学)

此词于 1959 年首次出现在美国。1961 年,在美国佐治亚工程大学召开的信息学会议上,提出了信息学的几个定义。泰劳尔(Talor)指出:“信息学包括了在交流领域中,从信息产生到最终使用的各个环节的所有职能。”海斯(Hayes)则是从信息产生系统的观点来认识,他说:“信息学是一门研究特定信息系统中信息传递、选择、处理和分析过程的学科。”1967 年,美国文献学会主席 L. 杰鲁尔(Jeroool)在给会员的一封信中写道:“信息学是研究信息的特性和活动、管理传递过程的手段以及保证信息的最有效利用所必需的信息处理技术。信息学研究范围包括:利用自然的和人工的系统提供信息的方法,利用号码有效进行信息的传递、存贮和检索,

研究信息处理设备和方法,如计算机及其程序设计系统。”他提出的定义,经会员略作修改后,被美国文献学会(ADI)采纳。1968年,该会改名为美国情报科学学会(ASIS)。《美国文献工作》(American Documentation)1970年改名为《美国情报科学学会会刊》(Journal of ASIS)。从此,信息学在美国扎下根,并不断完善和发展。

在前苏联,与 Information Science 对应的术语是 информатика,按米哈依诺夫(А. И. Михайлов)的解释,信息学是“研究科学信息的结构和基本性质,并研究科学交流所有过程的一般规律性的一门科学。”

在德国,人们采用 Informations und Dokumentationswissenschaft(信息与文献科学)一词表示信息学。在日本被译为情报科学。在我国大陆被译为情报科学、信息学,在台湾则被译为资讯学或资料科学。

2. Informatics(信息学或信息自动化处理技术)

此词最早出现在法国,法文为 Informatique。是1962年由 Philippe Dreyfus 工程师创造的新词。它由 INFORMATION 加 autoMATIQUE 构成,表示信息处理自动化的含义,尔后,在报刊杂志上使用的频率较高。同年苏联科学院院士 A. A. 纳尔凯维奇译成俄语,称为 Информатика,组词法相同。1966年法兰西科学院予以承认,定义为:“一门在技术、经济和社会领域里,用计算机处理被视为人类知识载体的信息的一门学科和技术的总体。”因而,“Informatique”在汉语中译为信息处理技术更为确切。在日本被称为“情报工学”。

3. Information Sciences(信息科学)

这一术语是指和信息处理与系统有关的所有科学,即所有的信息(包括人类信息和非人类信息,生物信息和非生物信息,语义和非语义信息)研究的整合。美国著名学者马克卢普(F. Muchlup)编写的《信息研究:学科之间的通讯》一书认为信息科学应包括:哲学中有关信息范畴的研究,控制论和系统论中有关信息一般功能的研究,认识论,认知科学,大脑理论中有关人类精神信息的研究,符号学、语义学、象征学、图书馆学、文献学、档案学、博物馆学、传播学、新闻学、教育学等,信息科学的边缘学科,如信息社会学、信息经济学、信息管理学、信息心理学等,还有人工智能以及信息技术所包含的各种技术领域中有有关的研究。由此可见,信息科学是一门包罗所有信息研究的综合学科。信息科学是研究信息现象及其运动规律和应用方面的综合性学科,是以“三论”为理论主体,以电子计算机等为主要工具的一门新兴的横断学科。

申农创立的信息论是以通信系统模型为研究对象,以概率论和数理统计为工具,从量的方面描述信息的传输和提取方面的问题,特别是深入研究了通信和控制系统中普遍存在着的信息沟通与传播的共同规律,以及如何达到信息获取、变换、传递的最佳效果。与此同时,控制论的基本理论则由维纳在他的名著《控制论:或动物和机器中控制和通信的科学》一书中所奠定。维纳根据信息论的基本原理,进一步研究了机器和生物中信息传输、变换,处理和控制的一般规律,对于信息理论的发展做出了独特的贡献。系统论的早期发展是与生物学联系在一起的。第二次世界大战后,美籍奥地利生物学家贝塔朗菲(L. von Bertalanffy)创立的一般系统论开始崭露头角。系统论把事物作为一个整体和系统来研究,运用数学模型和类比同构等概念去描述系统的结构和行为,提出了适用于一切系统的模式、原则和规律。到了70年代,系统论的思想获得了新的补充和重大发展,这就是比利时物理学家普里高津等人提出的耗散结构理论和德国科学家哈肯提出的协同理论。这两种理论都着重于研究系统如何由无序状态转化为有序状态,探索这种转化条件、可能性、规律性和相应的机制,触及了系统理论的核心问题。信息论、

控制论、系统论在许多基本概念、基本思想、基本方法等方面有着共同之点,三者从不同的角度和侧面提出了解决共同问题的方法和原则。从信息角度看,信息论主要研究信息的认识问题;控制论主要研究如何利用信息来达到有目的的控制;系统论则主要研究如何利用信息来实现系统的优化。从这个意义上讲,信息论可认为是控制论和系统论的基础;控制论和系统论又从不同角度丰富和发展了信息理论。由此,“三论”互相融合、互相补充,共同奠定了信息科学的理论基础。

4. Library and Information Science(图书馆与信息学)

这一术语表示图书馆学与信息学是两个平行的学科,彼此相互联系,又各自独立。美国各大学相关学院、科系的名称多取此意义,与计算机科学是有差异的。我国也大体相同,国家学位评审委员会于1996年10月25日下发的关于授予博士、硕士学位的学科、专业目录中将图书馆学、情报学与档案学归入文献信息管理列于管理学一级学科。在德国和西欧其他一些国家则采用 Information and Documentation Science(信息和文献学)的称谓。这说明图书、文献、信息部门都是以知识信息作为研究对象的。在信息技术应用方面也几乎相互重叠,向信息概念靠拢。

(二)信息学的定义

如信息的概念没有统一的结论一样,信息学的定义在国内外学者中也未形成一致的意见。

前苏联学者米哈依洛夫指出,信息学是“研究科学信息的结构和基本性质,并研究科学交流所有过程一般规律的一门科学。”简而言之,他认为信息学是关于科学交流和科学信息的科学。信息学的产生是由于科学信息活动,是从科学研究活动和科学组织活动中分离出来而形成的相对独立的领域。因此,信息学既非图书馆学,也不是来自图书馆学母体,同样也非计算机科学。但米哈依洛夫在定义信息学概念时,将其研究对象仅局限于科学信息和科学交流过程,也存在着明显的局限性。

著名学者 B. C. 布鲁克斯主张,“信息的任务是探索和组织客观知识”。他从哲学的角度探索信息学基础理论,提出信息作用于知识的方程式,并主张活化静态的知识结构,提出把文献蕴含的知识绘成“知识地图”。其学术观点鲜明并具有新意。但他采用了波普尔“世界3”的多元论哲学观点,提出的知识结构也有不成熟之处。另一位美国著名的情报学专家,客观知识学派的代表人物 F. W. 兰开斯特(Lancaster)教授认为,信息学研究信息技术应用于知识交流的基本原理和规律,以及服务于知识交流的信息系统的设计原理。

1979年ISO(国际标准化组织)提出,“信息学是对信息的功能、结构、传递的研究和信息系统管理的研究”。我国标准(GB4894—85)《情报与文献工作词汇基本术语》把信息学定义为“研究情报获取、传递与使用的理论、规律与方法以及情报系统管理的学科”。

从术语学的观点出发,信息学研究人类社会中部分社会信息——以知识形态为主的,也包括新闻等非知识形态的那部分信息,阐明社会信息现象、信息过程和信息行为。而信息科学或广义信息学则研究一切信息——人类的、非人类的,生物的、非生物的,包括哲学、控制论和系统论中有关信息的研究。

由此可见,信息学是社会信息化的产物,其研究对象是社会信息现象、信息过程和信息行为,探索社会信息化的规律性。也就是说,信息学是研究信息获取、整序、传递和利用的规律性的一门科学。

(三)信息学研究内容

信息学主要研究以下几方面的内容:

1. 基本概念。如信息的概念、属性、特征以及与科技、经济、社会的关系;信息学研究对象、学科性质及与相关学科的关系。
2. 基本原理和理论体系。如研究基本概念之间的关系及组合形式;研究文献计量学的 6 个基本定律。
3. 信息产生源的研究。包括客观知识、知识性信息;信息序化的研究,涉及知识、信息从混沌到序化的过程和规律;信息传递的研究,涉及传播、交流理论,信息技术的应用以及信息价值转化方式和过程的研究。
4. 信息增值的研究。涉及信息用户、信息需求的类型、特点与规律的研究,信息转化为生产力的理论与实践的研究。
5. 信息系统的设计、运行与评价。
6. 信息的组织与管理。如信息政策、信息体系、运行机制、信息立法以及信息教育等。
7. 信息技术应用问题。计算机、通信、存贮技术等。

第三节 信息学基本原理及理论体系

一、信息学基本原理

纵观信息学的研究范畴,概括信息学的客观现象和基本规律,确认基本原理如下:

(一)信息产生原理

这一原理有比较成熟的理论基础,即文献计量学的 6 个定律(将在本书第二章详细介绍):

1. 作者分布定律:文献统计中的洛特卡定律(Lotka's law),是一个平方反比律。它的作用在于指示作者与文献数量的关系,描述科学生产率问题,即科技工作者论著的数量。

洛特卡定律可表达为 $f(x) = \frac{C}{x^2}$ 。

其中, C 为常数,论文数 x 与作者频率 $f(x)$ 之间存在着一定关系。这就是科学生产的平方反比率。

2. 文献分散定律:英国人布拉德福(S. C. Bradford)进行了大量的统计调查,于 1934 年发表了《特定主题的信息源》,首次用数学模式描述了文献分散的客观规律。其公式为:

$$R(n) = K/n(an+1)$$

式中: $R(n)$ ——对应于 n 的相关论文累积数;

n ——期刊等级排列的序号(级);

a ——第一级期刊中相关论文数 $R(1)$,也就是载文率最高的期刊中相关文献数;

K ——参数,等于分散曲线中直线部分的斜率,可用实验方法求得。当 N (等级排列的期刊总数)足够大时, $K \approx N$ 。

布氏定律及其区域分析法、图像观察法以及 3 种假设推理方法,对信息检索、信息搜集和选择核心期刊及核心文献都有实际应用价值。后来,布鲁克斯又从数学上对布氏定律的实际应用技术进行了探讨和完善。

3. 词频分布定律:美国语言学家齐普夫(G. K. Zipf)经过大量的统计,试图证明自然语言词汇的分布服从一个简单的定律,并称之为“省力原则”。表示为 $r^m \cdot f = C$ 。

齐氏定律在标引和词表编制理论及计算机信息检索中应用比较广泛。

4. 文献增长规律:美国科学史专家普赖斯(D. Price)在其著名的《巴比伦以来的科学》一书中,以科学杂志和学术论文为知识发展的两个重要标志为研究对象,对知识量的增长率进行了推算。通过曲线分析,发现科技文献增长与时间呈指数函数关系。如果用 $F(t)$ 表示时间 t 的文献量,则指数定律为:

$$F(t) = ae^{bt}.$$

a 是统计初始文献量, e 是自然对数底, b 是时间常数(持续增长率)。

考虑到某些因素,如物质的、经济的、智力的因素对文献增长率的影响,又有学者提出文献逻辑曲线增长理论,其方程为:

$$F(t) = \frac{K}{1 + ae^{bt}}$$

该曲线称为逻辑曲线。

5. 文献老化定律:目前描述文献老化规律的数学公式一般采用 $C(t) = Ke^{-at}$ 。

式中: $C(t)$ 表示发表了 t 年的文献引用频率; K 是常数; e 是自然对数的底, a 是老化率。

R. 巴尔顿和 R. 凯普勒提出用下式描述老化规律:

$$Y = 1 - \left(\frac{a}{e^x} + \frac{b}{e^{2x}} \right)$$

式中: $a + b = 1$;

Y 为经过一定时间被利用的某一学科或专业全部文献的相对部分;

x 为时间,以 10 年为单位。

6. 文献聚类定律:涉及文献耦合(document coupling)、同被引(co-citation)、自引、引用文献(citing document)、被引文献(cited document)等概念。

引证分析法用来确定核心文献(含期刊),一般采用引用次数法、被引次数法、引用被引并用法、引用被引比较法,同时还要考虑影响因数法,文献耦合原理以及综合法(被引、被摘及流通)等方法。

(二)信息序化原理

知识序化理论和方法是信息学研究的核心。在信息产生、序化、传递、增殖(吸收、利用)这 4 个主要现象中,序化起主导作用,也是最本质的现象。序化就是将杂乱无章、随机零散的知识进行整序、分析和综合成为人们利用的形态。事实上,人们认识和改造客观事物就是一种从混沌到有序、从初级有序到高级有序的进化过程。

信息序化原理是依据普利高津(I. Pringogine)的耗散结构理论确立的。耗散结构理论同样是人类信息现象和行为的基本原理。普利高津假设,一个系统形成耗散结构至少要有 4 个条件:①系统必须是开放系统;②系统必须处于远离平衡状态;③系统内部各要素之间存在着非线性的相互作用;④涨落导致有序。

人的知识系统是一个开放系统,可以表示为熵的变化: $ds = des + dis$

其中, des 为系统与外界交换的熵流, dis 为系统内部产生的熵流,来源于“知识单元”数目的增多和衰减等。为了维持和提高知识系统的功能,要求 $ds < 0$, 且 $|des| > |dis|$, 即由外界向系统输入足够的负熵流(知识信息),系统才趋于有序状态。

哈肯的协同论认为:各种不同系统在一定条件下,其内部各子系统之间通过非线性相互作

用产生协同效应,从混沌状态走向有序,从低级有序走向高级有序,或者从有序转化为无序。

此外,在信息检索理论方面,人们提出了一些基本原理和方法:

检索效率:J. W. 佩里(Perry)在 1955 年首先提出了查全率(R)和查准率(P)。1957~1968 年间,英国学者克勒维当(C. Cleverdon)经过反复实验证明发现 R 和 P 之间的互逆相关关系。另一学者斯维兹(J. A. Swets)提出: R 和 P 并不是理想的检索效率评价指标。检索效率综合评价指标应为覆盖率、查全率、响应时间、用户费力程度、输出格式等。此外,不少学者也试图用数学方法如集合论、决策论、非线性优化方法、模糊数学等应用于检索理论,而且近年取得了可喜的成果。

此外,信息检索理论还包括检索系统模式和检索语言的理论和方法。

(三)信息传递原理

研究信息传递交流的行为和过程。其主要研究领域有:信息传递、交流的模式,信息存取的理论和方法;研究正式和非正式交流的理论与方法;研究通信技术、联机检索系统、网络技术、接口技术、机器翻译、数字、图像、声音的传递技术及其应用等。

信息传递的先决条件应当是信息必须处于激发状态,即 $I \geq I_0$ 。信息传递的时间 T 与其自身的价值 I 和信息用户对信息的需求强度 F 成正比,与传递环绕阻力 f 成反比。其数学表达式为: $T = K \cdot \frac{I}{f} \cdot F$ 。

在信息传递理论研究方面,诸如申农在通信数学原理的研究成果,维纳的控制论,米哈依洛夫对交流理论的研究等,为信息传递理论和方法研究奠定了理论基础。近十几年来,在此领域又有很多新的突破。

(四)信息增殖原理

信息增殖原理即信息的吸收利用原理。

B. C. 布鲁克斯研究了信息对读者的作用,提出了描述信息作用于知识结构的基本方程式:

$$K[S] + \Delta I = K[S + \Delta S]$$

式中: $K[S]$ ——吸收信息前知识系统的状态;

$K[S + \Delta S]$ ——吸收信息后知识系统的状态;

ΔI ——该知识系统吸收的信息。

上述方程的一般解释是:知识结构 $K[S]$,由于吸收了 ΔI ,增长为 $K[S + \Delta S]$,新知识的增加是吸收了信息的结果。这一基本方程表达式对信息增殖原理作了精确的描述。

应该说,信息学不单是一门单纯研究“物”(文献)的学问,也不单纯是研究信息技术应用的学科,而是一门以人和社会为主体研究对象的学科。因此,要研究人的信息吸收,人的信息心理、信息行为及人吸收利用知识的过程应是信息学研究的主要内容。相应地,认知科学和知识工程便成为信息学的重点基础学科,来完善和发展信息学理论体系。

二、信息学理论体系

信息学理论体系是指信息学概念、原理及一切信息活动的内在联系和科学规律的理论综合的学科整体。在信息学理论建设上,除了研究信息学本学科的基础理论和基本方法,建立和完善理论体系外,还要注意应用研究及与分支学科和相关学科的关系,保持信息学理论体系的

整体性、系统性和科学性。

(一)信息学理论体系研究的现状

一门学科理论体系的建立和发展一般要经历经验阶段或理论体系孕育形成阶段,理论体系建立和稳定阶段,理论体系成熟阶段及理论体系预测等四个阶段。由于信息学这门新兴学科的理论研究落后于实践,有关信息学的研究对象的理论界定尚不统一,并且没有属于本学科的研究方法,当前信息学的理论体系的建立正处于从第一阶段向第二阶段的过渡之中。即其基本理论、分支理论,应用理论和技术方法有待于进一步成熟,仍需要实践的考验。因此,信息学作为一门学科体系还相当年轻。

1. 国外信息学理论体系研究

由于各国的历史、文化背景不同,信息实践方式、信息工作模式各异,对信息学理论的研究也各不相同,从而形成不同的理论体系。例如,60~70年代,以前苏联米哈依洛夫为代表的科学交流型的理论体系和以加菲尔德为代表的引文分析型体系。70年代末以德国的克劳斯为代表的综合理论体系;80年代产生了以英国布鲁克斯为代表的知识利用型理论体系;70~80年代兴起的以美国萨尔顿和日本北川敏南为代表的信息技术应用型体系以及以美国兰开斯特为代表的联机检索型体系;在东方则有印度拉德扎贡帕兰的信息使用体系。这些理论体系从不同研究角度出发而各具特色。

2. 国内信息学理论体系的研究

我国信息学理论体系的研究起步较晚。80年代以后,国内信息界的学者在学习国外信息学理论的基础上,也开始对信息学的理论体系进行研究,并发表了很多论文和著作。代表性的类型有以下几种:

①二元分类体系(即广义、狭义信息论);

②多元层次体系(即:理论—应用—技术三个层次;宏观—中观—微观;表层—中层—深层)。

③信息流螺旋环型体系(以信息产生、组织、传递、吸收4个环节以及与之相对应的基本原理为研究对象)。

④信息流特型体系(以研究问题出发,提出信息学研究规范。包括信息源——S规范;传递过程——P规范,后改为C规范;信息用户——U规范)。

⑤立体空间结构型体系(视信息学为受若干基本参量限制而产生的维度群组成的三维空间:基础科学、工程结构、研究结构、研究规范)。

⑥思维科学激活型体系(钱学森提出的理论体系)。这一理论体系创造性地提出了信息学的学科性质、学科特点等问题。

第四节 信息的量与质

一、信息量与不确定性

信息论创始人申农于1948年提出信息量的概念,所谓信息量是对一件事或一件东西的确定程度的度量。信息量的大小取决于表现信息内容的消息的不确定程度。换句话说,信息与不确定性有密切的关系:获取信息,则会在一定程度上消除不确定性,使不确定性变为确定的或较为确定的。因此,不确定性的变化程度就能反映所获取信息的多少,只要能够量度不确定性