

21

世纪普通高等教育规划教材



河南省“十二五”普通高等教育规划教材

信息管理学

(第二版)

—— 郭秋萍 主编 赵 静 任红娟 副主编 ——



化学工业出版社

21

世纪普通高等教育规划教材



河南省“十二五”普通高等教育规划教材

信息管理学

(第二版)

—— 郭秋萍 主编 赵 静 任红娟 副主编 ——



化学工业出版社

· 北 京 ·

《信息管理学》从信息管理理论和实践两个角度出发,介绍了信息管理理论的基本问题、相关信息技术、应用领域以及最新发展动态,使学生既能宏观地了解人类社会信息管理活动的客观规律,掌握信息管理的基本理论和方法,自觉运用所学的知识和技能丰富信息管理的实践活动,又能使学生通过把握信息源生命周期,通过微观、中观、宏观的信息管理三层理论学习,把理论与实践紧密地结合起来。另外,为了将最新的信息管理技术与技术发展的前沿相匹配,本书还介绍了数据挖掘技术、云计算技术和大数据技术相关内容。

《信息管理学》可作为高等院校信息管理类专业(信息管理与信息系统、图书馆学、情报学、档案学、公共管理等)、企业管理类和其他相关专业的本科生和研究生的教材或参考书,也可作为相关领域的科技工作者或各类企业的参考书。

图书在版编目(CIP)数据

信息管理学/郭秋萍主编. —2版. —北京:化学工业出版社, 2017.8

21世纪普通高等教育规划教材 河南省“十二五”普通高等教育规划教材

ISBN 978-7-122-29949-9

I. ①信… II. ①郭… III. ①信息管理-高等学校-教材 IV. ①G203

中国版本图书馆 CIP 数据核字 (2017) 第 136751 号

责任编辑:王淑燕
责任校对:边涛

装帧设计:张辉

出版发行:化学工业出版社(北京市东城区青年湖南街13号 邮政编码100011)

印装:北京云浩印刷有限责任公司

787mm×1092mm 1/16 印张17 字数431千字 2017年9月北京第2版第1次印刷

购书咨询:010-64518888(传真:010-64519686) 售后服务:010-64518899

网 址: <http://www.cip.com.cn>

凡购买本书,如有缺损质量问题,本社销售中心负责调换。

定 价: 45.00 元

版权所有 违者必究

前 言

人类要在现实世界中生存与发展,就需要物质、能量、信息三大要素,其中信息是最具活力、最具影响的重要资源,它对物质、能量具有驾驭作用,支配着物质、能量的流动方向。学会获取信息、存储信息、处理信息和传递信息,已经成为现代人必备的基本技能之一;从事信息管理的科学和服务,也将是当今最为热门和尖端的学科和活动。信息管理伴随着人类社会的产生而发展,信息管理学的形成和发展是多学科交叉研究的成果,是一门建立在数学、管理科学、传播学、信息科学与信息技术的基础上,融合多门学科、多个领域的知识,涉及微观信息管理、中观信息管理及宏观信息管理等几个层面。

信息管理学是信息管理与信息系统专业的一门重要的专业基础导论课程。该课程的目的是向学生介绍本专业领域的基础性和概论性知识,为学习后续课程提供支持。随着信息管理学的发展,新的理论和技术不断涌现,同时,经过几年的教学实践发现教材存在的一些不足,需要更新教材内容。

在第二版中,为了更加明晰信息管理学中相关概念的含义及其相互关系,引入钟义信教授对信息、知识、情报等相关概念的解析和辨析,丰富了信息管理学的理论支撑。为了将最新的信息管理技术与技术发展的前沿相匹配,增加了数据挖掘技术、云计算技术和大数据技术相关内容。根据信息管理学科当前的研究重点和前沿对相应章节进行了知识更新和补充,加入了信息资源质量评价等新内容,力求形成信息管理课程的学习和教学提供一站式的资源服务体系。

本书由郭秋萍主编,赵静、任红娟担任副主编。全书共分7章,第1和4.3章由赵静执笔,2.1、2.2、7.2~7.4和7.6由李伟超执笔,2.3由王海燕执笔,2.4节和2.5由周湘贞执笔,第3章由郭秋萍与任红娟共同执笔,4.1、4.2、4.4、4.5由武利红与郭亚军共同执笔,第5章与7.1由任红娟执笔,第6章与7.5由莫祖英执笔。在具体编写过程中,研究生吴婷婷和王艺枫同学收集整理资料,对教材做了通校工作。

本书内容丰富、体系完整、观点新颖,叙述深入浅出,穿插、辅助以最新、全面的信息管理应用案例,还附有参考文献,每章后均有帮助读者理解和思考的习题,在强调理论和方法重要性的同时,突出了实践性和可操作性。本教材有配套使用教学课件和相关教学资料,可以辅助教学使用。

由于时间仓促,作者水平有限,难免存在疏漏与不当之处,敬请广大同行及读者不吝赐教并给予宝贵意见,以便在修订中及时更正。

编者

2017年5月

目 录

第1章 信息管理学概论	1
1.1 信息	1
1.1.1 信息定义的谱系	1
1.1.2 信息的定义	5
1.1.3 与信息相关的几个概念	7
1.1.4 信息的类型	9
1.1.5 信息的特征	11
1.1.6 信息的度量	12
1.1.7 信息的功能	15
1.2 信息管理	16
1.2.1 信息管理的概念	16
1.2.2 信息管理的分类	17
1.2.3 信息管理的特征	22
1.2.4 信息管理的发展历程	23
1.3 信息管理学	26
1.3.1 信息管理学的产生背景	26
1.3.2 信息管理学的研究对象	27
1.3.3 信息管理学的理论体系	31
1.3.4 信息管理学的理论流派	32
1.4 信息管理学的学科基础	33
1.4.1 信息科学	34
1.4.2 管理学	37
本章小结	43
思考题	44
第2章 信息管理的技術基础	45
2.1 信息技术概述	45
2.1.1 信息技术的发展状况	45
2.1.2 信息技术的内涵	45
2.1.3 信息管理与信息技术的关系	46
2.1.4 信息技术在信息管理领域的应用阶段	47
2.2 数据挖掘技术	49
2.2.1 数据挖掘的由来	49

2.2.2	数据挖掘的定义	49
2.2.3	数据挖掘的处理对象	50
2.2.4	数据挖掘可发现的知识类型	51
2.2.5	数据挖掘的一般流程	55
2.2.6	案例	57
2.3	云计算技术	58
2.3.1	云计算的概念	58
2.3.2	云计算的体系结构	59
2.3.3	云计算的优势	60
2.4	大数据技术	62
2.4.1	新摩尔定律	62
2.4.2	大数据定义	63
2.4.3	大数据特征	63
2.4.4	大数据分析层次	63
2.4.5	大数据的相关开源组件简介	64
2.5	网络信息检索技术	65
	本章小结	68
	思考题	68
第3章	信息源及其分布规律	70
3.1	信息源概述	70
3.2	信息源的种类	71
3.3	信息源的分布规律	72
3.3.1	文献信息的增长规律	72
3.3.2	文献信息的老化律	73
3.3.3	文献信息的分散规律——布拉德福定律	76
3.3.4	著者分布规律——洛特卡定律	77
3.3.5	词频分布规律——齐普夫定律	78
3.3.6	马太效应	79
3.3.7	长尾理论	80
3.4	引文分析	80
3.4.1	引文分析的基本内容和方法	81
3.4.2	引文分析的测度及特点	82
3.4.3	学术文献的双引分析	84
3.5	网络信息计量学	87
3.5.1	网络信息计量学的产生与发展	87
3.5.2	网络信息计量学的研究对象	89
3.5.3	网络信息计量学的主要研究方法	90
3.5.4	网络信息计量学的主要特征及研究意义	92

本章小结	92
思考题	92
第4章 微观信息管理	93
4.1 信息搜集与组织	93
4.1.1 信息搜集	93
4.1.2 信息组织	101
4.2 信息存储与检索	110
4.2.1 信息存储	110
4.2.2 信息检索	113
4.3 信息交流	119
4.3.1 信息交流的内涵	119
4.3.2 常见的信息交流方式	119
4.3.3 常见的信息交流模式	120
4.3.4 社会信息交流	123
4.3.5 信息交流的影响因素	124
4.4 信息分析与预测	125
4.4.1 信息分析	126
4.4.2 信息预测	130
4.5 信息用户与服务	133
4.5.1 信息用户	133
4.5.2 信息服务	138
本章小结	142
思考题	142
第5章 中观信息管理	143
5.1 信息系统管理	143
5.1.1 信息系统管理概述	143
5.1.2 信息系统的规划	144
5.1.3 信息系统的开发	151
5.1.4 信息系统的维护和评价	158
5.2 信息市场管理	161
5.2.1 信息市场概述	161
5.2.2 信息商品	164
5.2.3 信息市场管理	166
本章小结	167
思考题	167
第6章 宏观信息管理	169
6.1 信息产业	169
6.1.1 信息产业概述	169

6.1.2	信息产业的特征及发展演进	173
6.1.3	信息产业的分类及测度方法	176
6.1.4	我国信息产业现状及发展趋势	179
6.2	信息政策与信息法律	181
6.2.1	信息政策与信息法律概述	182
6.2.2	信息政策及其体系	184
6.2.3	信息法律及其体系	191
6.2.4	我国国家信息政策、法律建设现状分析	196
6.3	信息教育与信息伦理	201
6.3.1	信息教育	201
6.3.2	信息伦理	202
本章小结		205
思考题		206
第7章 信息资源管理与应用		207
7.1	信息资源管理概述	207
7.1.1	信息资源管理的概念	207
7.1.2	信息资源的类型和特点	208
7.1.3	信息资源管理的手段	210
7.1.4	信息资源管理的目标和任务	211
7.1.5	信息资源管理的过程	212
7.1.6	信息资源管理的功能	213
7.1.7	信息资源的配置	214
7.2	社会信息化	220
7.2.1	社会信息化的层次	220
7.2.2	社会信息化的构成要素及发展过程	221
7.3	企业信息资源管理	222
7.3.1	企业信息化	222
7.3.2	企业信息资源管理的内容、规划及组织结构	223
7.3.3	企业信息资源管理系统及其技术	227
7.3.4	中小企业信息资源管理	228
7.3.5	企业知识管理战略	231
7.4	政府信息资源管理	233
7.4.1	政府信息化	233
7.4.2	电子政务与政府信息资源管理	235
7.4.3	政府信息资源的开发与利用	237
7.4.4	政府信息公开	239
7.5	个人信息管理	241
7.5.1	个人信息管理理论	241

7.5.2 E-learning 与个人信息管理	242
7.5.3 构建个人信息管理的必要性	244
7.5.4 个人信息管理系统	245
7.6 信息资源质量管理	253
7.6.1 信息质量	253
7.6.2 信息资源质量管理的内容	256
本章小结	258
思考题	258
参考文献	259

第 1 章 信息管理学概论

随着现代科学技术特别是信息技术的发展,信息的数量越来越多、增长速度越来越快,出现了“信息爆炸”的现象;但人们在拥有大量信息的同时,却面临着“信息丰富,知识贫乏”的困境。上述问题解决的主要方法就是加强信息管理研究,对信息进行科学、有效的管理,使得人们可以快速查找到所需的信息,并得以有效利用。

1.1 信 息

信息一词可以说是当今社会使用最频繁的一个词,它出现在人们生活、工作、学习的方方面面,在各个领域都可以看到这个词的踪影。实际上,信息一词在我国古代就常有出现。唐朝诗人杜牧在《寄远》中有“塞外音书无信息,道傍车马起尘埃”;南唐诗人李中在《碧云集·暮春怀故人》一诗中有这样的语句“梦断美人沉信息,目穿长路倚楼台”;宋朝的李清照亦有“不乞隋珠与和璧,只乞乡关新信息”;至清朝,《红楼梦》第 16 回里,讲到贾政突然奉旨入朝,全府惶惶不安,后有赖大禀道,“小的们只在临敬门外伺候,里头的信息一概不能得知”。

信息对应的英文单词为“Information”,来源于拉丁文“Information”,法语中也是“Information”。在我国的港澳台地区,通常称信息为“资讯”,而日文中称为“情报”。

1.1.1 信息定义的谱系

人们很早就知道通过语言、动作等传递信息。结绳记事是文字发明前,人们所使用的一种存储信息和传递信息的方法,在一条绳子上打结,用以记事,大事大结,小事小结。神秘的八卦也是人类信息活动的结晶。在古代,信息及信息活动已广泛存在,不过信息一词还是一个普通词汇,不是科学术语,主要指消息、音讯。

最早把信息作为科学对象加以研究的是在通信领域,同时它也在系统论和控制论中得到了充分的关注。

(1) 哈特莱

“信息”最早作为科学概念被研究是在哈特莱(R. V. L. Hartley, 1888—1970) 1928 年撰写的《信息传输》一文中。他认为信息就是“(在通信符号表中)选择通信符号的方式”,且用“选择的自由度”来计量这种信息的大小。如,一个人在符号表中选择了“I am well”这样的符号,他就发出了“我平安”的信息;如果他选择了“I am sick”这样的符号,他就发出了“我病了”的信息。同时,不管符号所代表的意义是什么,只要从符号表中选择的符号数目一定,发信者所能发出的信息的数量就被限定了。若发信者只能从由 0 和 1 两个符号组成的符号表中选择符号且长度限定为 1,则发信者只有两种选择,即 0 和 1,这时能传达的信息量很小;若放宽限制,比如长度不超过 n ($n > 1$),则发信者可以选择的符

号串有 $2n$ 个,能传达的信息量就可以很大(取决于 n 的大小)。即,若符号表中符号的个数为 S ,被选符号序列的长度为 N ,则能表示的信息量的大小为 $I=N\log S$ 。哈特莱的思想和研究成果对后来申农的思想有很大的影响,为信息论的创立奠定了基础,提供了思路。

但是哈特莱的研究也有一定的局限性,主要表现在以下两点:一是只考虑了通信符号的选择方式,而没有考虑所选择符号的内容与价值意义;二是符号的选择者必然使信息具有个性色彩。

(2) 吉布斯

吉布斯(Josiah Willard Gibbs, 1839—1903)是美国物理化学家、数学物理学家,1871年任耶鲁学院数学物理学教授,是全美这一学科的第一个教授。

吉布斯首先将统计学引进物理学,使物理学不得不考虑事件的不确定性和偶然性。从而使人类在科学把握信息的意义上迈出了第一步。他指出熵是关于一个物理系统信息不足的量度,但因为技术条件的限制,他没有提出信息的数量理论。

(3) 申农

申农(Claude E. Shannon, 1916—2001,也有译作:香农、仙农)是数位通信理论和信息论的创始人,是影响人类社会进程的科学家之一。

信息作为一门严密的科学,主要应归功于贝尔实验室的申农。20世纪中期,申农从不同方面完成了关于信息的统计理论,推导出了信息测度的数学公式,他指出,信息是消除不确定性的东西,信息消除的不确定性的的大小就是信息量。1948年,年仅32岁的申农发表了信息论的奠基之作——《通信的数学理论》,从通信的角度研究了信息传递与度量问题,标志着信息论的诞生。

(4) 维纳

维纳(Norbert Wiener, 1894—1964),美国数学家,控制论的创始人。维纳在其50年的科学生涯中,先后涉足哲学、数学、物理学和工程学,最后转向生物学,在各个领域中都取得了丰硕成果,称得上是20世纪多才多艺和学识渊博的科学巨人。他一生发表论文240多篇、著作14本。他的主要著作有《控制论》(1948)、《维纳选集》(1964)和《维纳数学论文集》(1980)。

控制论的奠基人维纳把信息的概念引入了控制论,将信息的概念与人的认识、动物的感知活动联系了起来。他认为:“信息这个名称的内容就是我们对外界进行调节并使我们的调节为外界所了解时而与外界交换来的东西。”维纳的定义包含了信息内容与价值层面的意义。但其局限是:人与外界交换的不仅仅是信息,还有物质与能量。

维纳同法国物理学家布里渊一样,认为信息的实质是负熵:“正如熵是无组织(无序)程度的度量一样,消息集合所包含的信息就是组织(有序)程度的度量。事实上,完全可以将消息所包含的信息解释为负熵。”

维纳在《控制论》中指出:“信息就是信息,既不是物质也不是能量。”维纳在这里强调了信息的特殊意义。那么信息与物质、能量之间的关系到底是什么呢?

① 信息与物质、能量之间的区别。信息是物质的基本属性而不是物质本身,它可以独立存在,又不影响物质的存在与运动,它所表现的主要是物质运动的状态与方式。如,一个物体在运动,它的运动状态和状态改变方式等可以被摄影机拍摄下来,经过处理可以将其重新表现出来。此时,产生这种运动状态和方式的那个物体已经离开观察者,但它的信息已经被记录并保留下来。保留下来的仅仅是信息,而不是源物体本身。物质具有质量且遵守守恒定律,而信息本身没有质量且不遵守守恒定律;信息可以共享,而物质不存在真正意义上的

共享。

信息是物质的运动状态与方式，而能量则是物质做功的本领；能量转换遵循守恒定律，而信息转换不存在守恒现象；能量不能共享，而信息可以共享；能量为人类提供动力，而信息为人类提供知识与智慧。

② 信息与物质、能量之间的联系。信息与物质、能量之间的联系集中体现在三者都统一于物质，物质是第一性的，能量和信息都源于物质。物质是信息存在的基础，任何信息都不可能脱离物质而存在；任何物质的运动过程同时也是信息运动的过程，而任何信息运动的过程都离不开物质的运动过程。能量是信息运动的动力，信息的获取和传递都离不开能量，能量的传递与驾驭也离不开信息。信息与物质、能量还可以互相转化。信息虽然既不是物质也不是能量，但在一定条件下，信息可以转化为物质和能量。“知识就是力量”“知识就是生产力”等所表述的正是这种转化关系。

(5) 贝塔朗菲 贝塔朗菲 (Bertalanffy, 1901—1972)，美籍奥地利生物学家，一般系统论和理论生物学创始人，20 世纪 50 年代提出抗体系统论以及生物学和物理学中的系统论。贝塔朗菲系统论以一般系统为研究对象，一般系统都具有以下特点：第一，系统具有整体性；第二，系统由相互作用和相互依存的要素所组成；第三，系统受环境影响和干扰，和环境相互发生作用。而系统以上三个基本特点都和信息密切相关，系统之间各要素之间的相互联系是以信息的传递为基础的，系统与周围环境发生相互作用的过程中信息的交流也是不可避免的，由此系统才能保持其最核心的特点——整体性。

申农的信息论、维纳的控制论及贝塔朗菲的系统论，被认为是 20 世纪 40 年代创立并发展迅猛的系统理论的分支学科，虽然它们仅有半个多世纪，但是在系统科学领域中已是资深望重的元老，被称为“老三论”“旧三论”。

与“旧三论”相对应，20 世纪 70 年代以来陆续确立并获得极快进展的系统理论分支学科——耗散结构论、协同论、突变论，被称为“新三论”。它们虽然时间不长，但已经是系统科学领域中年少有为的成员。

① 耗散结构论。1969 年比利时物理学家、化学家普里高津提出了耗散结构论。它主要讨论一个系统内从混沌向有序转化的机理、条件和规律。一个远离平衡态的非线性的开放系统（不管是物理的、化学的、生物的乃至社会的、经济的系统）通过不断地与外界交换物质和能量，在系统内部某个参量的变化达到一定的阈值时，通过涨落，系统可能发生突变即非平衡相变，由原来的混沌无序状态转变为一种在时间上、空间上或功能上的有序状态。这种在远离平衡的非线性区形成的新的、稳定的宏观有序结构，由于需要不断与外界交换物质或能量才能维持，因此称之为“耗散结构”。

② 协同论。它亦称协同学或协和学，由德国物理学家赫尔曼·哈肯于 1977 年创立。协同论最根本的思想和方法是系统自主地、自发地通过子系统的相互作用而产生的系统规则。竞争与合作的方法是它的重要研究内容，协同学最基本的概念也是竞争与协作。

③ 突变论。法国数学家雷内托姆于 1972 年发表的《结构稳定性和形态发生学》一书中阐述了突变理论。它是研究客观世界非连续性突然变化现象的一门新兴学科，它用形象而精确的数学模型来描述和预测事物的连续性中断的质变过程。

信息的概念非常广泛，从不同的角度对信息可下不同的定义。在众多中外学者的观点中本书列出一些较为权威的：

① 信息是用以消除随机不定性的东西（申农，1948）；

② 信息就是信息,既不是物质也不是能量(维纳,1961);

③ 信息是事物之间的差异(G. Longo, 1975);

④ 信息是集合的变异度(Ashby, 1956);

⑤ 信息是系统的复杂性(张学文等);

⑥ 信息不是物质,它是物质状态的映射(张学文等);

⑦ 信息是使概率分布发生变动的东西(Tribes et. al., 1971);

⑧ 信息是负熵(布里渊,1956);

⑨ 信息是被反映的差异;

⑩ 信息是被反映的变异度。

迄今为止,科学文献中对信息的定义已经超过上百种,且仍呈增长趋势。为什么会有如此众多的信息定义?为什么信息科学发展到今天,信息的概念仍未达成共识?首先,信息具有普遍性,普遍存在于自然界、人类社会和思维活动中,人类社会发展到信息社会,信息已经渗透到社会的各个角落,信息问题日渐显化,逐渐变得普遍且意义重大,吸引了各个领域众多研究者的注意。其次,各行各业、各学科领域的学者具有不同的知识背景、不同的观察角度、不同的研究目的,因而在定义信息时就会受到各自知识背景、观察角度、研究目标的限制,约束和制约着人们从各自背景和角度认识信息、定义信息。这些“知识背景、观察角度、研究目标”被统称为信息定义所遵循的约束条件。依据不同的约束条件定义信息,就会有不同的信息理解,就会得出不同的信息定义。显然,从每一种观察角度所得到的信息定义能在某种角度反映信息的某种本质属性,把从所有观察角度得到的信息定义综合起来才能得到信息的全部本质属性。

由此可见,产生不同信息定义的根本原因,是各种定义所依据的约束条件不同。约束条件的多样性是导致信息定义出现“仁者见仁,智者见智”情况的根本原因。依据相同的约束条件,就会得出相同或相近的信息定义。若按照某种准则将这些各不相同的约束条件有序排列,就可以得到各种信息定义的有序排列。那么这种准则又是什么呢?由于信息定义约束条件的基本作用是约束了信息定义的内涵,限制信息的适用范围,因此,一个比较恰当的信息定义约束条件排序准则,可以是约束条件的“约束强度”。如果信息定义约束条件的约束强度比较弱,那么在这种条件约束下所得到的信息定义的内涵就会具有比较普遍的性质,定义的适用范围也就较为广泛;如果信息定义约束条件的约束强度比较强,那么在这种约束条件约束下所得到的信息定义的内涵就会具有比较特殊的性质,定义的适用范围就会比较狭窄。这样,按照信息定义约束条件的“约束强度”准则,运用系统分析方法,就可以将各种信息定义进行排序,钟义信在《高等人工智能原理》一书中提出:“将约束强度最弱,即最具普遍性的内涵和最具广泛性的适用范围的信息定义,排在最前面;将约束强度最强,即最具体特殊性的内涵和最具特定性的适用范围的信息定义排在最后面,就可以得到一个有序的‘信息定义的谱系’”。

钟义信最早提出约束条件下的信息定义谱系,指出信息定义约束条件有以下几种基本情形。

(1) 约束条件强度为零:本体论信息

若定义信息时无任何约束条件限制,得到本体论层次的信息定义。这是最基本也是最重要的信息定义。

(2) 约束条件强度最弱:认识论信息

强度最弱的约束条件(也是最有意义的约束条件)就是“存在认识主体”。若定义信息

时只存在这唯一的约束条件：必须从主体的立场出发来定义信息，这样得到的即为认识论层次的信息定义。为了保证“条件约束强度最弱”，主体必须是最具一般性的正常主体，即具有正常的观察能力、理解能力和判断能力，而不能有其他附加条件。显然，认识论信息的内涵不具有本体论信息内涵那样的普遍性，也没有本体论信息那样广泛的适用范围。有关认识论信息的概念将在下面介绍。

(3) 约束条件强度非最弱：各种特殊的认识论信息

若对认识论层次信息定义的约束条件（主体）增加更强的约束，就得到内涵更为特殊、使用范围更加狭窄、层次更低的信息定义，称为某种特殊的认识论信息。

(4) 约束条件强度最强：最特殊的认识论信息

从逻辑上说，若对认识论层次信息定义加上一切可能的约束条件，在完备的约束条件下就会得到内涵最为特殊、适用范围最狭窄、层次最低的信息定义，称为某种最特殊的认识论信息。

由此可得如表 1.1 所示的信息定义谱系，从上到下信息定义的层次依次下降，由普遍性依次过渡为特殊性。

表 1.1 表明：最高层次的信息定义没有任何约束条件限制，称为“本体论信息”，本体论信息的适用范围最广。在此基础上引入“站在主体立场上来定义信息”的最弱约束条件，最高层次的本体论信息定义就退化为次高层次的认识论信息定义，其内涵就更为特殊，适用范围更为狭窄。引入的条件越多，定义的层次越低，内涵越特殊，适用范围越窄。根据引入约束条件的强弱程度能够得到一系列不同内涵、不同适用范围的信息定义，构成信息定义的谱系。

表 1.1 信息定义谱系			
约束条件	定义名称	所属层次	适用范围
无	本体论信息	最高	最广
最弱	认识论信息	次高	次广
.....			
较强	较特殊的认识论信息	较低	较窄
.....			
最强	最特殊的认识论信息	最低	最窄

表 1.1 的信息定义谱系中，各个层次的信息定义不是相互独立，而是根据约束条件的减弱或增强相互转换。如本体论信息定义增加“主体”条件，就上升为认识论信息定义；将认识论信息定义减去“主体”条件，就会下降为本体论信息定义。其他各个层次的定义也是如此。因此，这是一个互相联系的有机体系。

在整个信息定义谱系中，本体论信息和认识论信息是最重要的。本体论信息是整个信息定义谱系的总根和源头，没有本体论信息，就没有其他层次的信息。认识论信息从认识主体与客体事物相互作用的角度揭示信息的性质，其他更低层次的信息都要从认识论信息演化而来。

1.1.2 信息的定义

关于“信息”的定义，本书引用钟义信教授的观点。钟义信在《信息科学原理》一书中指出：“信息是事物的存在方式或运动状态，以及这种存在方式或运动状态的直接或间接的

表述”。这里的“事物”泛指存在于人类社会、思维活动和自然界中一切可能的对象。“存在方式”指事物的内部结构和外部联系。“运动状态”则是指事物在时间和空间变化所展示的特征、态势和规律。

从这个概念可以看出,信息是有不同层次的,它既有事物本身的存在方式或运动状态,同时也有人们认识了该事物之后对事物的存在方式或运动状态的表述,即本体论层次的信息和认识论层次的信息。

(1) 本体论信息

本体论信息是指事物所呈现(所表述)的运动状态及其变化方式。“事物”泛指一切可能的研究对象,既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。“运动状态”是指事物运动在空间上展现的性状和态势,“运动状态的变化方式”是指事物的运动状态随时间而变化的方式。

客观世界的一切事物都在运动,都有一定的运动状态和状态变化方式,因此,一切事物都在产生信息。这是本体论信息的绝对性和普遍性。而一切不同的事物都具有不同的运动状态和状态变化的方式,这又是本体论信息的相对性和特殊性。在本体论信息定义中,没有出现主体因素,因此,本体论信息是一种客观存在,不以主体存在与否为转移,无论有无主体,或者是否被主体感受到,都不影响它呈现的运动状态及其变化方式。这是本体论信息一个非常重要的特征。

(2) 认识论信息

主体关于某事物的认识论信息,是指主体所表述的该事物运动状态及其变化方式,包括运动状态及其变化的外在形式、内在含义和效用价值。

本体论信息与认识论信息之间既有本质上的联系,又有原则上的区别。本质上的联系是指两者都关注“事物的运动状态及其变化方式”。原则上的区别则表现在前者是“事物”本身所表述的运动状态及其变化方式,后者是“主体”所表述的事物运动状态及其变化方式。它们之间的这种本质联系与原则区别,使得它们既有各自明确的内涵,又能够通过定义的约束条件相互转换,引入主体条件本体论信息定义就转化为认识论信息定义;取消主体条件,认识论信息定义就转化为本体论信息定义。

引入了主体条件,认识论信息概念就具有了比本体论信息概念更为丰富的内涵。这是因为,首先,作为正常的认识主体,它具有感觉的能力,能够感觉到事物运动状态及其变化方式的外在形式;其次,它具有理解能力,能够理解事物运动状态及其变化方式的内在含义;再次,它具有目的性,能够判断事物运动状态及其变化方式对于主体目的而言的效用价值。而且,对于正常的认识主体,事物的运动状态及其变化方式的外在形式、内在含义和效用价值这三者之间是相互依存、相互制约、不可分割的。所以,在认识论信息层次研究信息问题时,必须同时考虑事物运动状态及其变化方式的形式、含义和效用三个因素。

实际上,主体只有在感知了事物的运动状态及其变化方式,理解了它的含义,判明其效用价值,才算真正掌握了这个事物的认识论信息,才能做出正确的判断和决策。把这样同时考虑事物运动状态及其变化方式的外在形式(语法信息)、内在含义(语义信息)和效用价值(语用信息)的认识论信息称为“全信息”。也就是说,认识论信息就是全信息,它是语法信息、语义信息、语用信息的三位一体。

语法信息,指主体所表述的事物运动状态及其变化方式的外在形式。语义信息,指主体所表述的事物运动状态及其变化方式的内在含义。语用信息,指主体所表述的事物运动状态及其变化方式对于主体目标而言的效用价值。

图 1.1 表示全信息的概念模型，可以较为直观地表述语法信息、语义信息和语用信息的概念以及它们之间的相互关系。

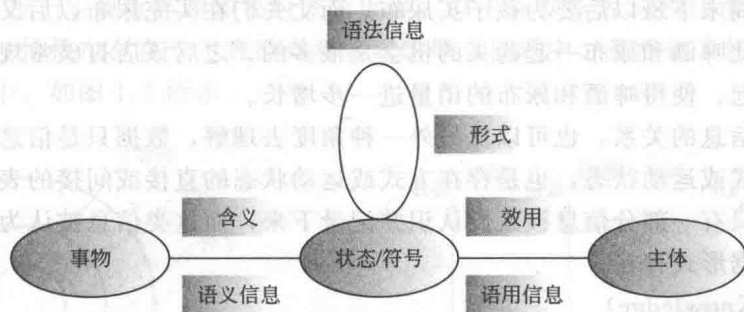


图 1.1 全信息的概念模型

① 当主体要表述本体论信息的时候，通常会用某种抽象符号及其序列来表示事物的运动状态及其变化方式。于是这个抽象符号及其序列就表述了这个事物的语法信息。

② 语法信息仅仅是一个抽象符号及其序列，没有任何含义，只有把抽象符号及其序列与具体事物关联起来时，抽象符号及其序列才有了内容，也就是语义信息。

③ 如果语法信息和语义信息与具体的主体关联起来时，那么该事物的运动状态及其变化方式的外在形式和内在含义对主体呈现出了效用价值，这就是语用信息。

④ 研究语义信息要以语法信息为基础，因为“含义”是针对具体的状态及其变化方式来说的。同样，研究语用信息要以语义信息和语法信息为基础，因为“效用”是针对具体的状态及其变化方式所具有的含义来说的。在这个意义上，基于语法信息和语义信息的语法信息具有“全信息”的含义。

语法信息是具体的，可通过认识主体的感觉器官直接感受到；语用信息是实在的，可通过认识主体实际体验体会到；但语义信息是抽象的，既不能通过感觉器官感受到，也不能通过体验体会到，只能在具体的语法信息和实在的语用信息基础上通过思维器官的逻辑演绎才能得到。

有了上述定义体系，有关信息的种种定义都可以在该体系中找到相应的位置。比如，维纳认为“信息就是信息，不是物质也不是能量”，对照本体论信息定义，“事物呈现的运动状态及其变化方式”虽源于事物，但又不是事物本身，更不是能量，可见，维纳的观点属于本体论信息。申农认为，“信息是能够用来消除不确定性的东西”，而只有“关于事物运动状态及其变化方式的表述”才能用来消除认识主体关于事物运动状态及其变化方式的随机不确定性。可见，申农的信息定义属于认识论信息的语法信息。总之，各种科学的信息定义都可在信息定义谱系中找到自己的位置。

1.1.3 与信息相关的几个概念

(1) 数据 (Data)

数据是关于自然、社会现象和科学实验的定量或定性的记录，而这些记录往往是零散的、片断的、没有联系的，而对其进行挖掘，找出其内在联系，形成有价值的、整体性的东西，称为信息。因此通常认为信息是数据经过加工处理后所得到的。如每张超市购物小票上面记载的就是杂乱的、没有联系的数据，而利用数据挖掘工具对这些数据进行分析就可以找出每个小票间的关系，而这种每个小票间的关系就是一种对数据加工处理之后产生的信息。“啤酒与尿布问题”是数据与信息间关系的最好例证。沃尔玛超市通过建立数据仓库，按周

期统计产品的销售信息,对这些数据进行数据挖掘。结果发现,每逢周末,位于某地区的沃尔玛连锁超市啤酒和尿布的销量很大。进一步调查表明,在美国有孩子的家庭中,太太经常嘱咐他们的丈夫周末下班以后要为孩子买尿布,而丈夫们在买完尿布以后又顺手带回了自已爱喝的啤酒,因此啤酒和尿布一起购买的机会是最多的。之后该店打破常规,将啤酒和尿布的货架放在了一起,使得啤酒和尿布的销量进一步增长。

关于数据和信息的关系,也可以从另外一种角度去理解,数据只是信息的一部分。信息是事物的存在方式或运动状态,也是存在方式或运动状态的直接或间接的表述。它是一个外延更大的概念,只有一部分信息被人们认识并记录下来,而这类信息被认为是数据,还有很多信息并不以数据形式存在。

(2) 知识 (Knowledge)

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与掌握,是人的大脑通过思维重新组合的、系统化的信息集合。它是对信息加工、吸收、提取和评价的产物,包括显性知识和隐性知识。通常被认为是“系统化、组织化的信息”。

(3) 情报 (Intelligence)

情报是对特定主体解决某个问题时有参考或决策价值的信息。它是特定的、有用的信息,通常具有其他类信息所不具备的保密性、难获得性的特点。情报具有三个基本属性:知识性、传递性和有效性。我国从日本引入情报这个词汇,而英语却依旧使用“Information”。现在我国已经全部改称“信息”,通常被认为是“有用的信息”。

(4) 消息 (Message, News)

通常把语言和声音、音乐、文字、符号、数据、图像等统称为消息,而消息中所给予接收者的新知识或消息中包含的有意义的内容是信息。人们在传递消息的过程中,实质是将其承载的内容传递给了接收者,因此消息是信息的外壳,信息是消息的内核。

(5) 信号 (Signal)

信号常应用于通信领域,描写与消息相对应的电信号或光信号,是传递信息的载体。信息在传递时,总要转化成一定形式的信号。同样的信息,可以用不同信号形式表现出来。所以说,信号是信息的载体,信息是信号的内容。

我国学者梁战平教授提出了著名的“信息链”理论,他认为“信息链”是由“事实”(Facts)→“数据”(Data)→“信息”(Information)→“知识”(Knowledge)→“情报”或“智能”(Intelligence)五个链环构成。简单地说,“事实”是人类思想和社会活动的客观映射;“数据”是事实的数字化、编码化、序列化、结构化;“信息”是数据在信息媒介上的映射;“知识”是对信息的加工、吸收、提取、评价的结果;“情报”“智能”则是运用知识的能力。换句话说,“事实”、“数据”、“信息”、“知识”、“情报”或“智能”五个链环组成“信息链”(Information Chain)。在“信息链”中,“信息”的下游是面向物理属性的,上游是面向认知属性的,作为中心链环的信息既有物理属性也有认知属性。

而对于数据、信息、知识和情报之间的关系,梁战平教授根据英国科学哲学家卡尔·波普尔(K. Popper)的“三个世界”的理论,提出以下观点。

① 并列关系。它们分别属于“信息链”中的一个要素。

② 转化关系。数据不会自动变成信息,信息也不会自动变成知识,数据、信息、知识同样也不会自动变成情报。实现从数据到情报的关键要素是人。是人通过信息组织与管理、知识组织与管理来实现信息、知识、情报的相互转化。知识本身也是一种信息,情报本身也是一种信息,相互之间可以转化。但是,知识、情报不是一般的信息,而是体现人的认知因