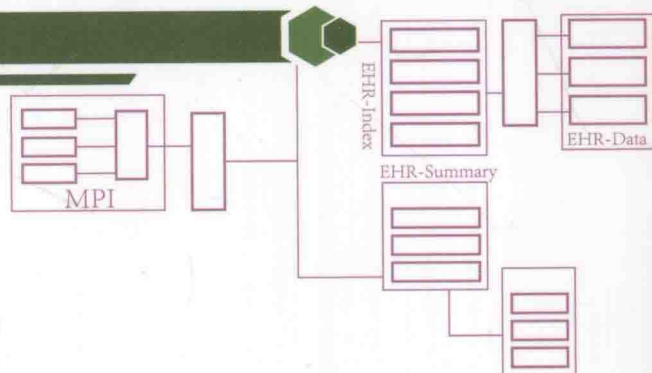


全国高等医药院校医药信息技术应用学科规划教材

医院信息系统分析与设计

施 诚 主 编



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

全国高等医药院校医药信息技术应用学科规划教材

医院信息系统分析与设计

施 诚 主编



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

“医院信息系统”是目前医学信息学领域内应用最多、发展最快的技术，也是目前医院信息化、数字化的支柱。

本书简要地介绍了医学信息学的发展概况，医院信息系统的基本概念，医院信息的采集与处理，医院信息系统的数据标准、功能特点等知识，并介绍了临床信息系统、电子病历、远程医疗及医学决策支持系统等内容，在此基础上进一步对医院信息系统的系统分析与系统设计过程做了详细的介绍，并以社区卫生服务系统和电子病历系统为实例进行了案例分析。本书共有6章，包括医院信息系统的基础知识、计算机信息处理技术基础、医院信息标准化、医院信息系统的特点、医院信息系统的分析与设计、医院信息系统开发案例。

本书既可作为高等医药院校信息类相关学科本科生学习医院信息系统设计与开发的教材，也可作为从事医院信息系统相关管理工作及开发、研究、运营等方面工作的人员的参考书。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

医院信息系统分析与设计 / 施诚主编. —北京：电子工业出版社，2014.1

全国高等医药院校医药信息技术应用学科规划教材

ISBN 978-7-121-22066-1

I. ①医… II. ①施… III. ①医院—管理信息系统—系统分析—医学院校—教材 ②医院—管理信息系统—系统设计—医学院校—教材 IV. ①R197.324

中国版本图书馆 CIP 数据核字（2013）第 289250 号

策划编辑：刘宪兰

责任编辑：谭丽莎

印 刷：三河市双峰印刷装订有限公司

装 订：三河市双峰印刷装订有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本：787×1092 1/16 印张：21.25 字数：472 千字

印 次：2014 年 1 月第 1 次印刷

定 价：42.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。



序

当今世界，社会与经济发展已经开始全方位进入信息化、数字化时代。伴随着物联网、大数据、云计算概念的出现，医药行业信息技术应用面临更广泛的需求和更高层次的要求。《2013年中国医卫行业信息化建设与IT应用趋势研究报告》显示，2012年中国医卫行业的IT投入达185.6亿元人民币，较2011年同比增长22.6%；2013年医卫行业信息化建设投入继续保持理性状态，呈现平稳增长趋势，IT投资规模约为225.5亿元人民币，较2012年同比增长21.5%。在大量资金投入的背景下，医药信息整合与有效利用的态势越来越明显。比如，医疗机构间的协同和信息交换、医院内部科室间的信息交换、医院和病人之间的信息交换等的需求越来越迫切；整合来自不同厂商的HIS、LIS、RIS、CIS、ERP等的数据库，以实现信息共享、流程协同变得越来越重要；医药企业产品的全过程质量控制、管理与质量信息可追溯要求和商业智能化应用，也给业界提出了医药信息技术应用水平提升的新要求。

为了适应当前国内外信息技术飞速发展的形势，满足业界对这些新理论、新技术、新方法的应用需求，同时也为了更好地贯彻落实教育部、财政部《关于实施高等学校本科教学质量与教学改革工程的意见》和教育部于2012年3月颁布的《教育信息化十年发展规划（2011—2020年）》的精神和要求，在电子工业出版社的积极支持下，2013年年初，全国19所高等医学院校的51位教师会聚南京，共同商讨编写一套“全国高等医药院校医药信息技术应用学科规划教材”，以加强全国高等医药院校医药信息技术应用学科的教材建设，满足新条件下对人才培养的需求。

“全国高等医药院校医药信息技术应用学科规划教材”共6册，分别为广东药学院周怡教授和新疆医科大学赵小龙教授共同主编的《医药信息分析与决策》、南京中医药大学施诚教授主编的《医院信息系统分析与设计》、南京中医药大学周金海教授主编的《医药物联网概论》、江西中医药大学章新友教授主编的《医药信息资源管理》、沈阳药科大学陈玉文教授主编的《医药企业资源计划》、广东医学院冯天亮教授主编的《数据

库原理及其医学应用》。在电子工业出版社及各参编单位领导的支持下，经过全体参编人员一年的努力，这套教材终于正式出版。

“全国高等医药院校医药信息技术应用学科规划教材”既可以作为高等医药院校医药信息技术应用学科本、专科学生和研究生相关课程的教材，也可以作为广大在职医务人员和护理人员进行信息技术相关继续教育的教材，同时也可以作为从事各类医学信息技术开发工作的技术人员的参考书。

我们希望这套教材的出版及使用，能够给中国医药行业信息化人才培养带来新的贡献，对高等医药院校医药信息技术应用学科教学水平的提高有所帮助。但是，医药行业信息技术发展速度很快，要求所有教材编写实时跟进变化，是不太容易做到的，加之编写时间及编者水平有限，新教材在内容编写上难免存在不足，甚至错误。在此，我们衷心希望在该套教材出版后，能够得到业内相关专家和读者的反馈意见，恳请使用本套教材的高校教师及相关人员能够不吝赐教，帮助作者继续努力提高教材编写质量，共同为高等医药院校医药信息技术应用学科教育的发展做出更大贡献。

中国医药信息学会 (CMIA)

医学信息学理论与教育专委会主任

施 诚

2014年1月1日



前言

随着计算机技术、网络技术及数据库技术的飞速发展,计算机信息处理技术已经逐渐深入到医院管理的各个方面。目前在医学信息学领域中,计算机应用最多、涉及面最广的是医院信息系统(Hospital Information System, HIS)。按照卫生部信息化工作领导小组办公室在《医院信息系统基本功能规范》中的定义,医院信息系统是指利用计算机软硬件技术、网络通信技术等现代化手段,对医院及其所属各部门的人流、物流、财流进行综合管理,对在医疗活动各阶段中产生的数据进行采集、存储、处理、提取、传输、汇总、加工生成各种信息,从而为医院的整体运行提供全面的、自动化的管理及各种服务的信息系统。由此可见,医院信息系统是建设信息化医院的支撑环境,国家卫生部对医院信息化工作高度重视。2002年,出台了《医院信息系统基本功能规范》。2009年4月,国务院发布了《关于深化医药卫生体制改革的意见》,其中明确提出要“建立实用共享的医药卫生信息系统”“以医院管理和电子病历为重点,推进医院信息化建设;利用网络信息技术,促进城市医院与社区卫生服务机构的合作”。2010年2月,卫生部在《关于公立医院改革试点的指导意见》中再次提出“以医院管理和电子病历为重点,推进公立医院信息化建设,提高管理和服务水平。研究制订医疗机构内部信息管理的规定和标准,充分利用现有资源逐步建立医院之间、上级医院和基层医疗卫生服务机构之间、医院和公共卫生机构、医保经办机构之间的互联互通机制,构建便捷、高效的医院信息平台”。目前,在我国的大、中城市,医院信息系统已经普及到社区医院。

医院信息系统内容丰富,涉及面广,它由很多具有独立功能的子系统构成,如门诊挂号系统、划价收费系统、医生工作站、护士工作站、药房管理系统、实验室信息系统、输血管理系统、医学影像存档与通信系统、手术室麻醉系统等。为了配合医院的正常工作,还有物资管理系统、设备管理系统、人事管理系统、财务管理系统与经济核算管理系统等。随着医院的管理理念的转变,管理特点从原先以财务经济为主体的管理方式转化为以病人为核心的管理方式,为了更好地为病人服务,出现了临床信息系统、电子病

历、病人咨询服务分系统、医疗统计系统、院长综合查询与分析系统,以及连接医院外部的远程医疗咨询系统、医疗保险接口、社区卫生服务接口等。总之,医院信息系统是建设现代化医院不可缺少的基础设施。

医院信息系统是一门综合了医学信息学、管理科学、信息科学、系统科学、计算机科学、网络通信技术及数据库技术的新兴的边缘学科。医药院校的学生今后要走上社会,进入医院或与医院相关的行业工作,因此,医药院校的学生学习一些“医学信息学”及医院信息化方面的基本知识,了解一些有关医院信息标准化、医院信息流程、医院信息系统的体系结构及功能特点等基本知识,以及了解一些有关建立医院信息系统的设计方法与技术是很必要的。

《医院信息系统分析与设计》共6章,第1章为医院信息系统的基础知识,主要介绍了数据、信息、知识、信息系统等的基本概念,医学信息学的基本知识及医院信息系统概况;第2章为计算机信息处理技术基础,主要介绍了信息在计算机内的表示、数制、医院数据采集与处理、计算机网络与数据库基本知识、数据仓库、数据挖掘、云计算、射频技术及移动平台等新技术的应用;第3章为医院信息标准化,主要介绍了医院信息标准化的基础知识、医院信息标准化的方法、国外医学信息标准化工作的概况、我国医疗卫生信息标准化工作的近况;第4章为医院信息系统的特点,主要介绍了医院管理信息系统的功能特点,临床信息系统、电子病历的功能与特点,远程医疗、医学决策支持系统、医院信息系统的安全管理等内容;第5章为医院信息系统的分析与设计,主要介绍了软件工程的基本知识、医院信息系统的可行性分析、医院信息系统的系统分析、医院信息系统的系统设计、医院信息系统的系统实施等内容;第6章为医院信息系统开发案例,该章通过对社区卫生服务系统和电子病历系统两个案例的分析与设计,系统地介绍了有关信息系统的开发过程的一些基本知识。

《医院信息系统分析与设计》是全国高等医药院校医药信息技术应用学科规划教材,既可作为高等医药院校信息类相关学科本科生学习医院信息系统设计与开发的教材,也可作为从事医院信息系统相关管理工作及开发、研究、运营等方面工作的人员的参考书。

本书在编写过程中得到了南京海泰信息技术有限公司和上海金仕达卫宁软件股份有限公司的大力支持,他们给本书提供了一些相关的图片资料,在此表示衷心的感谢!

施 诚

2013.12.18

目 录

第 1 章 医院信息系统的基础知识	1
1.1 信息的基础知识.....	2
1.1.1 信息、数据、知识与载体的概念.....	2
1.1.2 信息的特性与分类.....	5
1.1.3 信息的收集、分析与传输.....	10
1.1.4 信息的计算与信息的数字化.....	13
1.2 信息系统的基本概念.....	16
1.2.1 系统的定义与分类.....	16
1.2.2 信息系统的组成及功能.....	18
1.2.3 信息系统的结构与特点.....	20
1.2.4 与信息系统相关的理论.....	21
1.2.5 信息系统的基本原理.....	24
1.3 医学信息学概况.....	26
1.3.1 医学信息学概念及发展简史.....	26
1.3.2 医学信息的特点.....	28
1.3.3 数字医学概述.....	30
1.4 医院信息系统概述.....	32
1.4.1 医院信息管理的意义.....	32
1.4.2 医院的组织结构.....	33
1.4.3 医院信息系统的目标及特性.....	34
第 2 章 计算机信息处理技术基础	37
2.1 信息在计算机内的表示.....	38
2.1.1 计算机内部数据、媒体、数字媒体与多媒体的概念.....	38

2.1.2	数制	40
2.1.3	数值信息在计算机内的表示	43
2.1.4	文字信息在计算机内的表示	44
2.1.5	图形、图像信息在计算机内的表示	46
2.1.6	声音信息在计算机内的表示	48
2.1.7	视频信息在计算机内的表示	50
2.2	医院数据的采集与处理	51
2.2.1	医院数据采集概述	51
2.2.2	医院数据的分类与特点	52
2.2.3	医院数据的采集方法	54
2.2.4	医院数据的存储与管理	55
2.2.5	医院信息数据化的作用	58
2.3	计算机网络的基本知识	58
2.3.1	计算机网络概述	58
2.3.2	计算机网络的体系结构	61
2.3.3	数据通信技术	66
2.3.4	常用网络设备与协议	70
2.3.5	网络安全管理	73
2.4	数据库的基本知识	75
2.4.1	数据管理技术的发展	75
2.4.2	数据库系统的组成与特点	76
2.4.3	数据模型	77
2.4.4	数据库技术的发展	79
2.5	新技术	80
2.5.1	数据仓库	81
2.5.2	数据挖掘	83
2.5.3	云计算	90
2.5.4	知识库	93
2.5.5	射频技术及移动平台	95
第3章	医院信息标准化	99
3.1	医院信息标准化的基础知识	100
3.1.1	标准化的基础知识	100
3.1.2	信息标准化的相关概念及其作用	104
3.1.3	医院信息化及标准化	106
3.2	医院信息标准化的方法	113
3.2.1	数据元与元数据	113

3.2.2	信息分类的方法	115
3.2.3	信息编码的方法	118
3.3	国外医学信息标准化工作	126
3.3.1	国外医学信息标准化组织简介	127
3.3.2	国际疾病分类 ICD	128
3.3.3	人类与兽类医学系统术语 SNOMED	131
3.3.4	美国卫生信息传输标准 HL7	135
3.3.5	医学影像与传输协议 DICOM	137
3.3.6	CDA 临床文档结构	142
3.3.7	CCOW	143
3.3.8	IHE	143
3.3.9	LOINC	144
3.3.10	Arden syntax	145
3.4	我国医疗卫生信息标准化工作	145
3.4.1	概述	145
3.4.2	我国医院信息系统的功能、建设及评价规范的相关标准	147
3.4.3	卫生信息数据的相关标准	149
3.4.4	中医药信息有关标准	151
第 4 章	医院信息系统的特点	155
4.1	医院信息系统的功能特点	156
4.1.1	医院信息系统的功能划分	156
4.1.2	医院信息系统的功能概述	157
4.2	临床信息系统	157
4.2.1	临床信息系统概述	158
4.2.2	临床信息系统平台简介	159
4.3	电子病历的功能与特点	170
4.3.1	电子病历概述	170
4.3.2	电子病历系统的定义	171
4.3.3	电子病历系统的发展和评估	172
4.3.4	中医电子病历系统的特点	173
4.3.5	电子病历的数据存储与管理功能	175
4.4	远程医疗	176
4.4.1	远程医学概述	176
4.4.2	远程医学系统的组成及远程医疗技术的特点	180
4.4.3	远程医疗的几种应用模式	184
4.4.4	中医远程医疗	187

4.4.5	远程医疗的发展趋势	188
4.5	医学决策支持系统	189
4.5.1	医学决策支持的基本技术	190
4.5.2	医学决策支持系统的功能	191
4.6	医院信息系统的安全管理	193
4.6.1	临床信息系统安全风险评估	193
4.6.2	医院信息安全技术	195
4.6.3	防火墙	198
4.6.4	患者隐私保护功能	200
4.6.5	医院信息系统的灾难恢复体系	201
第5章	医院信息系统的分析与设计	205
5.1	软件工程的基本知识	206
5.1.1	软件工程概念	206
5.1.2	软件工程过程模型	206
5.2	医院信息系统的可行性分析	207
5.2.1	医院信息系统的规划	208
5.2.2	医院信息系统的开发方式与方法	210
5.2.3	医院信息系统的可行性研究与开发计划	212
5.3	医院信息系统的系统分析	214
5.3.1	系统分析概述	214
5.3.2	医院信息系统的需求分析	215
5.3.3	医院信息系统的组织机构和业务功能调查	217
5.3.4	医院信息系统的业务流程分析	220
5.3.5	医院信息系统的数据流程分析	222
5.3.6	医院信息系统的数据字典建立	224
5.3.7	医院信息系统的功能/数据分析	226
5.3.8	医院信息系统的分析报告	227
5.4	医院信息系统的系统设计	228
5.4.1	医院信息系统的总体设计	229
5.4.2	医院信息系统的代码设计	231
5.4.3	医院信息系统的输入、输出及界面设计	232
5.4.4	医院信息系统的处理过程设计	234
5.4.5	医院信息系统的数据库设计	235
5.4.6	医院信息系统的系统设计说明书	237
5.5	医院信息系统的系统实施	238
5.5.1	医院信息系统的物理系统实施	238

5.5.2 医院信息系统的测试及运行	239
5.5.3 医院信息系统的测试报告及操作使用说明书	242
第6章 医院信息系统开发案例	247
6.1 社区卫生服务系统的概要设计与案例分析	248
6.1.1 社区卫生服务系统概述	248
6.1.2 社区卫生服务系统的案例——用户需求分析	248
6.1.3 社区卫生服务系统案例——系统设计	250
6.1.4 社区卫生服务系统案例——数据存储	251
6.1.5 社区卫生服务系统案例——功能分析	252
6.2 电子病历系统的案例分析与设计	257
6.2.1 电子病历系统概述	257
6.2.2 电子病历系统案例——系统分析	257
6.2.3 电子病历系统案例——数据存储与交换	258
6.2.4 电子病历系统案例——功能分析	260
附录 A 中英文对照	275
附录 B 医院信息系统基本功能规范	287
参考文献	321

第1章

医院信息系统的基础知识

内容提要

本章介绍了信息的基础知识，信息系统的基本概念和基本原理，医学信息学及数字医学的概况，并对医院信息系统做了一些简要的概述。

1.1.1 信息、数据、知识等基本概念

关于“信息”的定义很多，《辞源》指出：“信息”一词，最早见于《说文解字》：“信，诚也。从人，从言。”《说文解字》：“言，直言也。从言，从口。”《说文解字》：“口，人发声之器也。从凵，从凵，从口。”《说文解字》：“凵，象形，象口开之形。”《说文解字》：“口，人发声之器也。从凵，从凵，从口。”

“信息”究竟是什么？以下介绍一些学者的观点：信息论的创始人香农(C.E.Shannon)对“信息”的定义是：“信息是用来消除不确定性的东西。”现代控制论的创始人维纳(W.N.Viener)对“信息”的定义是：“信息是减少不确定性的东西。”维纳的定义是：“信息是减少不确定性的东西。”维纳的定义是：“信息是减少不确定性的东西。”维纳的定义是：“信息是减少不确定性的东西。”维纳的定义是：“信息是减少不确定性的东西。”

1.1 信息的基础知识

在人类历史的长河中,随着生产工具和科学技术的发展,人们经历了狩猎时代、农业时代、工业时代,至今进入了信息时代。在信息时代中,“信息”这一词汇的使用频率很高,如“商品信息”、“物价信息”、“证券信息”、“交通信息”、“市场信息”、“经济信息”、“科技信息”、“医学信息”、“医院信息”、“临床信息”、“护理信息”、“卫生信息”、“信息技术”、“信息科学”、“信息高速公路”、“信息社会”等。人们可以通过书籍、报刊、信件、电话、电视等多种途径获取自己所需的信息;随着 Internet 的出现,人类社会进入了互联网时代,它把人与人之间交流信息的距离缩短了,人们可以通过 E-mail、网络电话、MSN、QQ、WWW 浏览器、微信等网络工具随时从网上获取信息、交换信息。Internet 促使人们可获取的信息成倍增加,使“信息”的增长呈爆炸趋势。

“信息”是现代社会的很重要的“要素”。在现代社会中,人们随时随地都在自觉地或不自觉地与信息打交道。随着社会的发展,信息所起的作用也越来越明显,例如,研究生在做开题报告前需要检索、查询大量与该课题相关的信息,以借鉴别人的经验,避免重复别人的工作;某药厂花几十万资金购买生产某种新药的专利;几名特工人员利用虚假信息可以颠覆一个国家,它的作用相当于数个师的作战部队的力量;一条有价值的商业信息可以使商家获得巨额利润;一个正确的临床诊断信息可以拯救一名患者的生命……在这里,我们看到了“信息”在人们的生产实践与社会实践中的重要性,“信息”像传统的“物质”和“能量”一样,已成为组成现代信息社会的一个很重要的“要素”,它正在改变人们的生活环境和生存方式。

1.1.1 信息、数据、知识与载体的概念

关于“信息”的定义很多,《牛津字典》指出:“信息,就是谈论的事情、新闻和知识”;《韦氏字典》指出:“信息,就是在观察或研究过程中获得的数据、新闻和知识”;“信息”的英语单词为 information,《新英汉词典》上将 information 解释为通知、消息、情报、知识及资料等;《日语广辞苑》记载:“信息是所观察事物的知识”;《现代汉语词典》记载:“信息即消息;信息论中指用符号传送的报道,报道的内容是接收符号者预先不知道的”。

“信息”究竟是什么?以下介绍一些学者对“信息”的定义。

信息论的创始人香农(C.E.Shannon)对“信息”做了如下定义:“信息是用来消除某种不确定性的东西。”现代控制论创始人维纳认为:“信息就是信息,不是物质,也不是能量。”他同时指出:“信息就是我们在适应外部世界,并且使这种适应反作用于外部世界的过程中,同外部世界进行相互交换的内容的名称。”钟义信教授在《信息科学与技术导论》一书中对信息的定义:“本体论层次的信息,就是事物运动的状态和(状态改变的)方式。认识论层次的信息,就是认识主体所感知或所表述的事物运动的状态和方式。”李宗荣教授在他的《医学信息学导论》一书中指出:“任何一个有目的的系统,

都必然是材料、能量和信息的和谐结合，材料构成系统的形成，能量产生运转的活力，信息是指挥系统动作的灵魂……信息是事物的属性及内在联系的表征。”

国际标准化组织 (International Organization for Standardization, ISO) 对“信息”下的定义是：“信息是对人有用的数据，这些数据将可能影响到人们的行为与决策。”ISO 对“数据”下的定义是：“数据是对事实、概念或指令的一种特殊的表达形式，这种特殊的表达形式可以用人工的方式或者用自动化的装置进行通信、翻译转换或者进行加工处理。”根据这一定义，通常意义下的数字、文字、图画、声音、动画、影像等都是数据，因为它们都能负载“信息”——“有用的数据”，它们均可以通过人工的方式（如计算机）进行处理。

综上所述，ISO 对“信息”所下的定义中已经覆盖了香农的内容，因为从“数据”到“信息”的过程，也就是通过处理数据消除了部分不确定性，使数据转化为有用信息的过程。它也覆盖了维纳的内容，即这些可能影响到人们的行为与决策的“数据”，也是人们适应外部世界，并且同外部世界进行相互交换的内容。事实上，ISO 对数据的定义也阐明了信息的本质。按照 ISO 对信息的定义，“信息”的基础是“数据”，“数据”是对事实、概念或指令的一种特殊的表达形式，这里的事实就是指客观的事物，因此从本质上来看，这种表达形式体现了客观事物的属性，以及客观事物与客观事物之间的内在联系。在本教材中，将按照 ISO 的定义来认识“数据”和“信息”。

在现代社会，“知识”正如“信息”一样，使用频率很高，如汽车知识、法律知识、军事知识、世界知识、中文知识、数学知识、化学知识、物理知识、卫生知识、健康知识、统计知识、农业知识、科普知识、股票知识、气象知识、共享知识、产销知识、知识产权、知识界、知识分子等。通过搜索引擎百度，在互联网上查询“知识”，可以查到约有 100 000 000 项符合查询要求的网页。

“知识”根据《现代汉语词典》和《现代汉语辞海》的解释，指人们在改造世界的实践中所获得的认识和经验的总和。根据《韦氏字典》1997 年的定义，“知识”是人们通过实践、研究、联系或调查获得的关于事物的事实和状态的认识，是对科学、艺术或技术的理解，是人类获得的关于真理和原理的认识的总和。也就是说，“知识”是人们在处理信息的基础上，所获得的对研究对象更深层次的认识。

数据、信息与知识之间具有何种关系呢？数据是信息的基础，信息是对人们有用的数据，即已经按照某种规则进行筛选过的数据，知识是在信息的基础上进一步提炼而成的关于自然和社会的认识和经验的总和，也是目前人们希望通过运用计算机技术在大量的数据中寻找的数据的内在关系，即指导人们从大量的数据中认识信息的某种规律。

当今人类正处于信息时代，随着信息技术的高速发展，人们积累的数据量急剧增长。在这样一个信息爆炸的时代，如何有效地管理这些数据，从数据的浩瀚海洋中及时发现有用的信息，提高信息的利用率；如何使数据能真正为人们的决策生成和战略预测服务都是人们迫切需要解决的难题。为此，一个新的研究方向——计算机数据挖掘和知识发现技术应运而生。1995 年，在加拿大蒙特利尔召开了第一届知识发现和数据挖掘国际

学术会议。1997年,创刊了该领域的第一本学术刊物 *Knowledge Discovery in Database*。此后,知识发现与数据挖掘的研究工作逐步成为数据处理领域的一个热点。

总之,数据是信息的源泉,信息是知识的基础。但是这些概念都是相对的,如一张化验报告,对化验室来讲是经过处理数据后获得的信息,对临床医生来讲是分析疾病的数据;同样,在知识挖掘的过程中,又将已经积累的许多知识视为数据。

目前,“数据”与“信息”的关系有点像“非典”与“SARS”的关系,虽然“非典”在国内几乎众所周知,但它毕竟并非科学的概念。“非典”是世界卫生组织(World Health Organization, WHO)提出的,2003年3月12日,WHO在历史上第一次通过新闻发布会发出全球警告,新闻发布会标题使用的名称是非典型肺炎(Atypical Pneumonia)。美国国立医学图书馆医学百科全书对于“非典”的定义为:“非典”是由肺炎军团菌及肺炎衣原体等细菌引起的。而新爆发的传染病由病毒引起,它的临床表现、传染途径、流行病学、治疗方法及预后和“非典”很不相同。3天以后,即3月15日。WHO在发出的全球紧急旅游劝告中首次采用SARS(Severe Acute Respiratory Syndrome)。SARS反映了这种疾病的严重性、急性、呼吸道发病三个特征与肺炎军团菌及肺炎衣原体等细菌引起的“非典”完全不一样。由于问题的严重性,WHO立即建立了由加拿大、法国、德国、中国及中国香港特别行政区、日本、荷兰、新加坡、英国和美国等10个实验室组成的合作网络。此后,SARS就成为国际上该疾病的通用名称。4个星期以后,即4月16日,WHO的新闻发布会宣布在人身从未见过的新的冠状病毒成员被确定为SARS的原因。但是在我国,由于“非典”先入为主,所以至今新闻媒体还将SARS称为“非典”。数据与信息的关系也有点类似于此,人们已经习惯于混淆这两个不同的概念。

“信息”是人们对数据进行加工处理后所获得的对人有用的数据。我们在这里也可以推导出“数据”既不是物质,也不是能量,而是“信息”的载体。

什么是载体呢?载体是承载信息的媒体。这里举一个通俗的例子,若我们将“数据”比作一部货车,“信息”比作货物,则货车是用来运载货物的,而数据肩负着传递信息的使命,因此叫作载体。

信息与货物的区别:货物是实体,是物质;信息是依附在实体上的某些对象的属性,它是抽象的。信息的传递要依赖物质,信息是不可能脱离物质而独立存在的。例如,信息在人们的大脑里,是以大脑为载体的;当人们把它说出来时,它就以声音为载体;把它写在纸上,它就以纸和文字为载体,若用电话传递,则它就以声音和电流为载体。

再举一个数据的例子,数据15W是一串数字加字母的符号,单纯的15W可以说没有什么确切的含义,但是它若出现在某一家用电器的外包装上,那么对于购买者来说,它就是一个重要的信息,即该家用电器的功率是15W;再如12℃、20\$、牛顿等,单纯的12℃、20\$、牛顿等是没有确切含义的,但是将“12℃”确定为今天的平均气温,将“20\$”确定为《医院信息系统教程》这本书的价格,将“牛顿”确定为力学中的一种度量单位,而不是一位科学家的人名时,这些符号就有了确切的含义,它们都转变为

对某些人有用的信息。

再如“脾”字，如前面提到的“牛顿”一样，它是一个符号，是数据，没有确切的含义。而中医和西医用不同规则对“脾”这个符号进行了解释，使它在各自领域中有了不同的含义。

中医认为“脾”属脏，位于中焦，在膈之下，和胃相表里。“脾”的主要功能包括主运化、主升清和统摄血液等。按照中医理论，“脾”和肌肉、四肢、口唇等相关，是人们的后天之本。这里的“脾”是中医所理解的功能性的、抽象的“脾”。

西医认为“脾”属于网状皮系统，位于腹腔的左上方，呈暗红色，质软而脆。脾脏在胚胎时期是一个重要的造血器官，人出生后它能产生淋巴细胞和单核细胞。脾脏内的巨噬细胞能将衰老的红细胞、血小板和退化的白细胞吞噬消灭。它还能吞噬血液中的细菌、原虫和异物。脾有丰富的血窦，可储存一定量（约200ml）的血液，在机体剧烈运动或爬山或突然失血时，脾的平滑肌收缩，放出血液以补充机体的需要。脾中的淋巴细胞还能制造抗体。这是西医所理解的具有解剖基础、生理基础的“脾”。

再如，临床医生需要获得患者的疾病诊断信息。他可以选用现有的各种载体，尽可能多地获取与诊断疾病相关的数据：既可以采用中医“望、闻、问、切”的传统方法，也可以采用测量体温、血压，血常规化验、肝功能化验，CT、核磁、B超、心电图、脑电图等多种手段。一般情况下，临床医生不会漫无边际地收集数据，往往通过运用他的经验或知识，有目的地、选择性地收集他所需的数据，然后对这些数据进行加工处理，最后获得与患者诊断结果相关的数据——“信息”。这里的体温、血压、化验数据、图像，以及中医的四诊数据等均是患者当时体征的反映，它们既不是物质，也不是能量，而是医生明确诊断信息所必需的数据。

总之，“信息”是人们通过对“数据”进行加工处理后所获得的有用的数据，“数据”既不是物质，也不是能量，而是“信息”的载体。在计算机中，所有的数字、文字、图画、声音、动画、影像等都是数据，都是可以加工成人们所需的“信息”的原材料。

1.1.2 信息的特性与分类

1. 信息的特性

信息既不是物质，也不是能量，信息就是信息。它既体现了事物的某种客观属性，也呈现了事物与事物之间的某种客观联系。因此，信息具有很多特性。下面对信息的特性做一些简单的介绍。

1) 信息不是物质，信息可以共享、可以复制

任何具体的物质，当它被移动到别处之后，原来的地方就不再存在这一物体了。而信息则不同，信息可以复制或传播给他人。例如，当某人将知识化的信息传递给他人后，他本人并没有丢失自己的知识，相反地，由于在传递过程中反复使用，知识反而更加巩固和充实。再如中央电视台转播的世界杯足球赛的信息，全国有几亿人在同时观看，大家都可以共享。共享性是信息与物质和能量的最大区别。