

“十二五”高等学校信息管理与
信息系统专业规划教材

国家社科基金项目（11BTQ019）阶段成果

信息计量学

XINXI JILIANG XUE

肖 明 等编著

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

清华大学出版社 中国高等院校计算机专业教材
清华大学出版社 中国高等院校计算机专业教材

清华大学出版社 中国高等院校计算机专业教材

信息计量学

清华大学出版社 中国高等院校计算机专业教材

清华大学出版社



清华大学出版社

“十二五”高等学校信息管理与信息系统专业规划教材
国家社科基金项目（11BTQ019）阶段成果

信息计量学

肖 明 等编著

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

信息计量学是在文献计量学的基础上发展起来的一门重要的新兴学科,也是当前国际学术界研究最活跃的专业领域之一。本书从理论、方法、应用三个角度构建了信息计量学的内容体系,全面、系统地讲解了信息计量学的理论基础、经典规律、分析方法、工具软件、前沿进展,还集中探讨了其在学科评价、大学评价、专利评价、科技预测等领域的具体应用。各章后面均附有练习题帮助学生巩固所学知识。本书还提供配套的 PPT 教学课件。

本书可作为高等院校信息管理与信息系统、管理科学与工程、信息资源管理、电子商务、情报学、图书馆学、档案学、出版发行学、科学学与科技管理、科学评价与预测等专业的信息计量学教材,也可作为广大信息工作者、图书情报与档案工作人员、科研人员、评价人员及管理业务的参考书。

图书在版编目(CIP)数据

信息计量学 / 肖明等编著. — 北京 :
中国铁道出版社, 2014. 8
“十二五”高等学校信息管理与信息系统专业规划教材
ISBN 978-7-113-18945-7

I. ①信… II. ①肖… III. ①文献计量学—高等学校
—教材 IV. ①G350

中国版本图书馆 CIP 数据核字(2014)第 155678 号

书 名: 信息计量学
作 者: 肖 明 等编著

策 划: 巨 凤
责任编辑: 周 欣 何 佳
编辑助理: 孙晨光
封面设计: 一克米工作室
责任校对: 汤淑梅
责任印制: 李 佳

读者热线: 400-668-0820

出版发行: 中国铁道出版社(100054, 北京市西城区右安门西街 8 号)
网 址: [http:// www.51eds.com](http://www.51eds.com)
印 刷: 三河市华业印务有限公司
版 次: 2014 年 8 月第 1 版 2014 年 8 月第 1 次印刷
开 本: 787 mm×1092 mm 1/16 印张: 21 字数: 509 千
书 号: ISBN 978-7-113-18945-7
定 价: 39.80 元

版权所有 侵权必究

凡购买铁道版图书, 如有印制质量问题, 请与本社教材图书营销部联系调换。电话: (010) 63550836
打击盗版举报电话: (010) 51873659

前言

Preface

信息计量学是在文献计量学的基础上发展起来的一门重要的新兴学科，也是当前国际学术界研究最活跃的专业领域之一。

本书参考学时数为 48 学时。由于各院校、各专业的课程在内容设计、课时要求、教学条件等方面不尽相同，所以建议使用本书的老师们从实际出发，在现有教学内容的基础上进行适当增减。

本书作为高等院校“信息计量学”主教材，是在课题组成员从事国家社科基金项目“基于多方法融合的中外图书馆学情报学知识图谱实证研究（项目编号：11BTQ019）”以及讲授“信息计量学”“网络计量与评价”等课程的教学、科研成果的基础上写成的。全书分为 8 章，从理论、方法、应用三个角度构建了信息计量学的内容体系，全面、系统地论述了信息计量学的理论基础、经典规律、分析方法、工具软件、前沿进展，还集中探讨了其在学科评价、大学评价、专利评价、科技预测等领域的具体应用。

本书已在北京师范大学进行过多轮次的教学试用，受到北京师范大学学生的普遍好评。作为一个多学科交叉的研究领域，本书具有以下特色：

（1）本书主要采用案例驱动编写方式，每章正文前有引例，引出与本章内容相关的问题。各章均提供本章指南、知识框图、知识扩展、名人名言、本章小结、习题、参考文献以及配套的 PPT 教学课件。

（2）理论联系实际。参加本教材编写的主要人员均为国内从事信息计量教学科研工作的高校或者研究所的教学、科研人员，不仅撰写了大量论著，而且主持了多项科研项目，具有丰富的教学科研经验。

（3）部分案例直接改编自相关科研项目，能够引领学生了解科研前沿，做到学以致用。

因此，本书可作为高等院校信息管理与信息系统、管理科学与工程、信息资源管理、电子商务、情报学、图书馆学、档案学、出版发行学、科学学与科技管理、科学评价与预测等专业的信息计量学教材，也可作为广大信息工作者、图书情报与档案工作人员、科研人员、评价人员及管理业务的业务参考书。

本书由肖明等编著，全书由肖明组织编写，本书编写人员主要是北京师范大学、中国科学技术信息研究所、中国科学院等单位中从事信息计量学教学和科研工作的相关老师，具体分工如下：第 1 章（张薰、肖明）、第 2 章（王卫）、第 3 章（佟贺丰、袁军鹏）、第 4 章（袁军鹏）、第 5 章（刘兰）、第 6 章（肖明）、第 7 章（黄彬、王卫、张薰、肖明）、第 8 章（袁军鹏）。全书最后由肖明进行统稿和修改，并对各章中的相关案例、知识扩展进行了增补和删改。本书在编写过程中参考了国内外有关知识

图谱、科学计量学、信息可视化的大量著作和文献，还查阅了大量网络资料，在此对所有原作者表示感谢！

由于编者水平有限、编写时间紧迫，不足和错误之处在所难免，恳请读者不吝批评指正，相关意见和建议请发送到邮箱：ming_xiao02@sohu.com。

编 者

2014年4月

目 录

Contents

第 1 章 信息计量学概述	1
1.1 信息计量学的产生与发展	5
1.1.1 信息计量学的由来	5
1.1.2 信息计量学的产生背景	6
1.1.3 信息计量学的发展简史	8
1.2 信息计量学的含义与特点	11
1.2.1 信息计量学的含义	11
1.2.2 信息计量学的特点	11
1.3 信息计量学的学科体系	12
1.3.1 信息计量学的研究目的和意义	12
1.3.2 信息计量学的研究对象	12
1.3.3 信息计量学的研究内容	13
1.3.4 信息计量学与其他学科的关系	16
1.4 信息计量学的研究现状与发展趋势	19
1.4.1 信息计量学的研究现状	19
1.4.2 信息计量学的发展趋势	22
本章小结	24
本章习题	24
第 2 章 信息计量学的数学基础	25
2.1 统计学概述	29
2.1.1 统计与统计学	29
2.1.2 统计研究	29
2.1.3 统计信息的计量	30
2.1.4 抽样	32
2.2 数据的收集、整理与描述	34
2.2.1 数据的收集	35
2.2.2 数据的整理	36
2.2.3 数据的描述	37
2.3 统计推断	39
2.3.1 参数估计	39

2.3.2	假设检验原理	40
2.3.3	参数检验	41
2.3.4	非参数检验	43
2.4	多变量统计分析	45
2.4.1	相关分析	45
2.4.2	回归分析	46
2.4.3	主成分分析	49
2.4.4	聚类分析	51
2.5	图论基础	52
2.5.1	图论概述	52
2.5.2	图的基本概念	53
2.5.3	图的计算机表示	56
2.5.4	路径与连通性	57
2.5.5	树与生成树	59
	本章小结	60
	本章习题	60
第3章	信息计量学的理论基础	63
3.1	信息计量学的信息学基础	66
3.1.1	信息论	66
3.1.2	熵、互信息与自信息	66
3.2	信息计量学的社会学基础	67
3.2.1	马太效应	67
3.2.2	人类行为与最省力法则	69
3.2.3	社会网络分析	70
3.3	信息计量学的情报学基础	72
3.3.1	知识网络	72
3.3.2	知识地图	76
3.3.3	信息可视化	78
3.4	信息计量学的网络科学基础	80
3.4.1	Web 理论	80
3.4.2	小世界理论	82
3.4.3	复杂网络	84
	本章小结	85
	本章习题	85
第4章	信息计量学的经典规律	87
4.1	信息增长规律	89
4.1.1	指数增长模型	90

4.1.2	逻辑增长模型.....	92
4.1.3	国内外信息增长规律研究的进展.....	93
4.1.4	信息增长规律应用举例.....	95
4.2	信息老化规律.....	97
4.2.1	信息老化规律概述.....	97
4.2.2	信息老化规律应用举例.....	101
4.3	布拉德福定律.....	102
4.3.1	布拉德福定律的基本内容.....	102
4.3.2	布拉德福定律的发展与趋势.....	107
4.3.3	布拉德福定律的前提条件与局限性.....	109
4.3.4	布拉德福定律应用举例.....	110
4.4	齐普夫定律.....	113
4.4.1	齐普夫定律的形成和确立.....	114
4.4.2	齐普夫定律的基本内容.....	115
4.4.3	齐普夫定律的前提与局限性.....	115
4.4.4	齐普夫定律应用举例.....	116
4.5	洛特卡定律.....	119
4.5.1	洛特卡定律的基本内容.....	119
4.5.2	洛特卡定律应用举例.....	122
	本章小结.....	124
	本章习题.....	124
第 5 章	信息计量学的数据来源.....	125
5.1	传统数据来源.....	131
5.1.1	图书.....	131
5.1.2	期刊.....	132
5.1.3	科技报告.....	132
5.1.4	会议文献.....	134
5.1.5	专利文献.....	135
5.1.6	标准文献.....	137
5.1.7	学位论文.....	138
5.1.8	产品资料.....	139
5.1.9	技术档案.....	139
5.1.10	报纸.....	140
5.2	引文数据来源.....	141
5.2.1	国外引文数据库.....	141
5.2.2	国内引文数据库.....	150
5.2.3	网络引文.....	158
	本章小结.....	162

本章习题	163
第 6 章 信息计量学的工具软件	165
6.1 统计分析软件	168
6.1.1 SPSS 软件	168
6.1.2 SAS 软件	180
6.1.3 MATLAB 软件	187
6.1.4 R 语言	191
6.2 信息可视化软件	196
6.2.1 Pajek 软件	196
6.2.2 Ucinet 软件	200
6.2.3 Bibexcel 软件	204
6.2.4 HistCite 软件	210
6.2.5 CiteSpace 软件	213
本章小结	223
本章习题	223
第 7 章 信息计量学的应用举例	225
7.1 期刊评价	230
7.1.1 期刊的评价方法与信息计量指标	230
7.1.2 基于信息计量的期刊评价原则与类型	233
7.1.3 基于信息计量的期刊评价实例	235
7.2 科研人才评价	236
7.2.1 基于信息计量的科研人才评价指标	237
7.2.2 基于信息计量的科研人才评价原则	239
7.2.3 基于信息计量的科研人才评价实例	241
7.3 学科评价	243
7.3.1 基于信息计量的学科评价的内容与类型	244
7.3.2 基于信息计量的学科评价指标	246
7.3.3 基于信息计量的学科评价原则	250
7.3.4 基于信息计量的学科评价实例	251
7.4 大学评价	257
7.4.1 信息计量在大学评价领域的应用与发展	257
7.4.2 基于信息计量的大学评价原则	259
7.4.3 基于信息计量的大学评价指标及方法应用	259
7.4.4 基于信息计量的大学评价实例	267
7.5 专利评价	269
7.5.1 信息计量在专利评价领域的应用与发展	269
7.5.2 基于信息计量的专利评价指标	269

7.5.3 基于信息计量的专利评价原则	270
7.5.4 基于信息计量的专利评价实例	271
7.6 科技预测	273
7.6.1 信息计量在科技预测领域的应用与发展	274
7.6.2 基于信息计量的科技预测原则	275
7.6.3 基于信息计量的科技预测指标及方法应用	276
7.6.4 基于信息计量的科技预测实例	278
本章小结	281
本章习题	281
第8章 信息计量学的研究前沿	283
8.1 h 指数	285
8.1.1 h 指数的概念	285
8.1.2 h 指数的优缺点	286
8.1.3 h 类指数	290
8.1.4 h 指数的应用	292
8.2 引文分析	293
8.2.1 引文分析概述	294
8.2.2 引文分析主要指标	300
8.2.3 引文分析法的应用	302
8.3 网络日志分析	306
8.3.1 网络日志分析概述	306
8.3.2 网络日志分析应用软件	308
8.4 网络链接分析	311
8.4.1 网络链接分析概述	311
8.4.2 网络链接分析的研究方向	313
8.5 文本挖掘	313
8.5.1 文本挖掘概述	314
8.5.2 自动分词原理与算法	315
8.5.3 文本特征提取原理与算法	317
8.5.4 文本分类基本思想	318
8.5.5 聚类算法基本思想	319
8.5.6 关联规则基本思想	321
本章小结	324
本章习题	324
参考文献	325



第 1 章

信息计量学概述



鲁索
1949—

罗纳德·鲁索(Ronald Rousseau), 数学和图书情报学双博士, 国际著名信息计量学专家, 被誉为信息计量学之父。曾获国际计量学领域最高奖——普赖斯(Price)奖。现担任ISSI委员会主席。是中国科学院文献中心、大连理工大学、浙江大学、上海大学客座教授, 河南师范大学荣誉教授, 同时还是Scientometrics(科学计量学)等諸多学术杂志的编委会成员, ISSI第九届科学计量学与信息计量学国际学术会议学术委员会主席。

信息计量学是信息科学中有关信息测度方面的科学, 也包含用应用数学的方法对信息现象的研究。



学习目标

信息计量学是情报学领域中一门重要的基础学科, 它是在社会信息化环境下, 伴随着情报学的定量化研究产生和不断发展的。本章学习目标是掌握信息计量学的产生、发展以及含义与特点, 同时对于信息计量学的研究内容、研究对象和研究目的有明确的认识, 更好地理解信息计量学与其他学科的关系; 在此基础上, 了解信息计量学的研究现状与研究热点, 并掌握信息计量学的发展趋势。本章重点在于了解信息计量学的发展简史以及理论基础, 而要全面理解信息计量学的研究热点和发展趋势有一定难度, 希望读者能够借助相关参考书更好地理解相关内容, 真正掌握信息计量学的基础知识。



知识框图



开篇案例

地图引导我们在地球上探索了几百年, 人们得以更好地在世界上生存。科学地图能以同样的方式, 使我们能够利用可视化成果对知识进行探索, 还能够帮助我们理解今天所产生的海量科研数据间的层次关系, 拥有独特而有趣的视角。从人们开始对文献进行定量研究到现在经历了将近一百年的时间, 在这将近一百年的时间内, 文献计量学由萌芽走向了成熟。然而, 在新时代的背景下, 随着网络和网络相关技术的发展, 信息计量学研究发生了巨大变化。例如, 随着信息分析技术尤其是可视化技术的不断发展, 利用

信息计量学的方法,通过研究活动中各要素的行为模式和规律,以及结构中各层次各类别元素的相互关系,利用 CiteSpace II 等信息可视化软件,可以绘制出相应的知识图谱,以可视化的图像直观地展现所研究的内容彼此之间的关系。

美国 Sandia 国家实验室计算、计算机与数学中心的 Klavans R 和 Katy Börner 等人绘制了一系列图谱,展示了该中心关于科学共同体、科学结构和科学范式等方面映射技术的成果,这些图谱可以帮助国家、企业以及研究人员找到自身在相关领域中的位置。图 1-1~图 1-5 是其中的一些图片精选。

(资料来源: <http://scimaps.org/atlas/part2.html>, 有修改)。

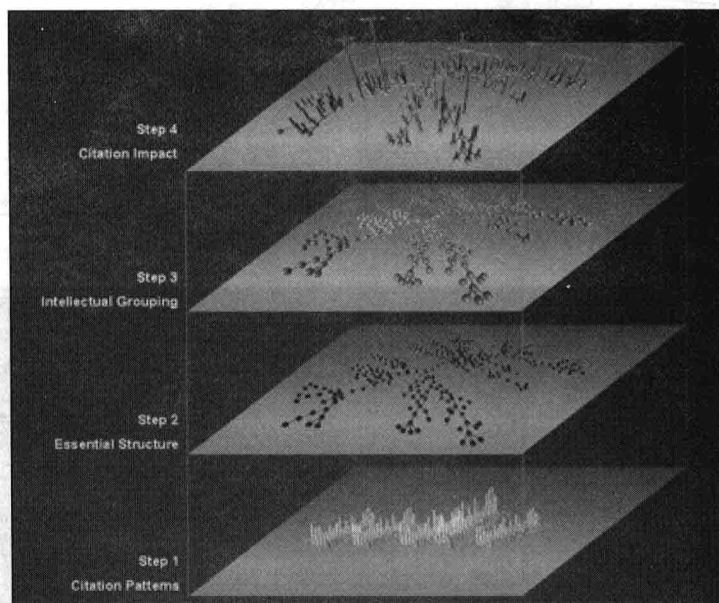


图 1-1 陈超美“科学领域知识结构的可视化”(2001)

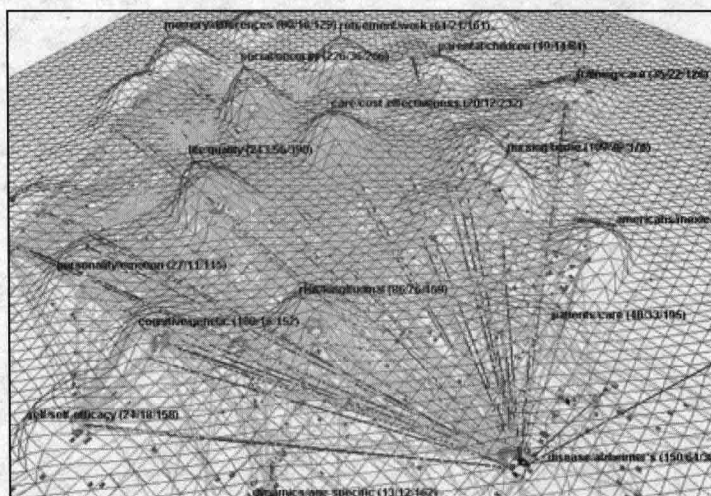


图 1-2 论文与资金的链接关系图(2003)

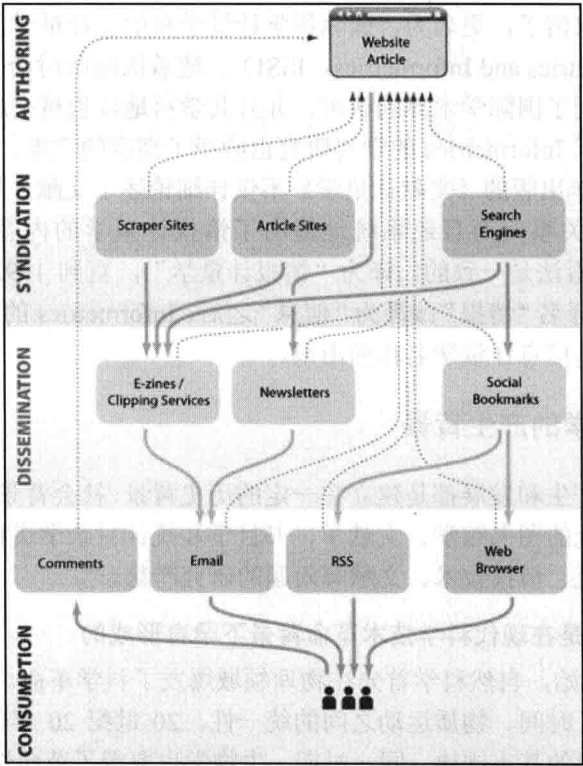


图 1-5 网络论文生命周期示意图

1.1 信息计量学的产生与发展

1.1.1 信息计量学的由来

信息计量学（原称情报计量学）名称最早出自德文 *Informetrie*，是由德国学者奥托·纳克（Otto Nacke）在 1979 年联邦德国的《文献工作通报》上首先提出，纳克等人后来继续从事这方面的研究，并于 1980 年 3 月在法兰克福市召开了第一次情报计量学讨论会，纳克教授在会上宣传了他提出的“情报计量学”术语。

对应的英文单词则为 *Informetrics*，我国将其译为“情报计量学”，后来改为“信息计量学”。刘廷元认为这个英文单词最早见于 1980 年美国科学基金会（NSF）公布的年度研究项目的标题中，也有人认为是日本杂志《情报管理》与苏联的《情报科学文摘》杂志将其转译为英文的。

而 *Informetrics* 被真正接受是从 1987 年在比利时的迪彭贝克（Diepenbeek）召开的“文献计量与信息检索理论”国际会议开始的，布鲁克斯（Brookes）当时建议将 *Informetrics* 包括在 1989 年即将在加拿大召开的会议名字中，这次会议因此被命名为“文献计量、科技计量和情报计量国际会议”。1991 年在印度的班加罗尔召开的第三届会议的名字是“国际计量学会议”，标志着这个术语的最终被接受。至 1995 年 6 月，在美国芝加哥召开的学术会议改名为“第五届科学计量学和信息计量学国际会议”，文献计量学被包含在内

而在会议名称中被取消了,更名为“国际科学计量学和信息计量学学会”(International Society for Scientometrics and Informetrics, ISSI),随着国际研讨会会议名称的变化,说明了情报计量学得到了国际学术界的认可,并且其学科地位也越来越突出。

我国学术界对于 Informetrics 的学科研究也给予了高度的重视。早在 1981 年就有相关论文发表。1988 年出版的《文献计量学》不仅详细论述了文献计量学、科学计量学和情报计量学之间的关系,而且还系统地提出了情报计量学的内容框架。我国学者对 Informetrics 的译名看法是一致的,译为“情报计量学”,直到 1992 年我国相关部门将 Information 对应的译名“情报”改译为“信息”之后,Informetrics 的中文译名也改为“信息计量学”,这就是信息计量学名称的由来。

1.1.2 信息计量学的产生背景

信息计量学的产生和发展都是建立在一定的历史背景、社会背景和学科背景之上的。概言之,它是对传统的图书馆学、文献学、书目学和统计目录学进行相应的变革,并适时地增加对科学交流、情报技术、文献等问题的研究产物。

1. 信息计量学是在现代科学技术革命背景下孕育形成的

从 19 世纪末开始,自然科学首先在物理领域爆发了科学革命。20 世纪初,相对论的创立揭示了空间、时间、物质运动之间的统一性。20 世纪 20 年代,量子力学体系成功地揭示了微观世界的基本规律。同一时期,生物学也取得了革命性的突破。20 世纪以后,以原子能工业、电子计算机、空间技术和遗传工程为标志的现代技术革命全面展开。当前,以信息技术为基础的信息资源网络化、电子化极大地推动了社会信息化的发展。

现代科学技术革命不仅从根本上改变了社会生产力的结构,提高了劳动生产率,引起了社会政治、经济和文化生活各方面的深刻变化,同时也对科学本身产生了深远的影响,形成了高度分化又高度综合的科学体系。科学活动日益社会化和科研规模的不断扩大,一方面为科学计量学的发展提供了一定的社会条件和动力,另一方面也促进了科学交流的蓬勃发展,使文献的数量急剧增加,形成了科学文献体系,成为科学交流系统的重要组成部分,为文献计量学的发展奠定了基础,同时也促进了科学的信息化过程。社会信息系统及其实践不断加强,提出了加强系统管理与信息管理的科学化、量化的要求,特别是信息资源网络化的发展为信息计量提供了超出文献计量、科学计量的更大空间,并已形成了网络信息计量学的基础,这些都是信息计量学形成和发展的重要基础。

2. 信息计量学是在文献计量学等相关学科的基础上拓展演变而成的

情报学形成之初,正遇上传统的统计书目学向文献计量学和科学计量学的转变时期。在这个转变时期,普理查德在 1969 年提出了用文献计量学取代传统的统计书目学,并认为文献计量学是“把数理和统计方法应用于研究图书馆及其交流中介”的学科;苏联学者纳里莫夫和穆利钦科在 1969 年提出了科学计量学概念,认为科学计量学是“研究分析作为情报(信息)过程的科学的定量方法”。因此,一些注重定量研究的早期情报学家很自然地就将文献计量学或者科学计量学看成是自己的耕耘园地,进而形成了文献计量学或科学计量学是“情报科学的特殊方法的综合”的观念。一些情报学家加入文献计量学研究队