

高等学校文科教材

# 情报语言学基础

增订二版

张琪玉 著

武汉大学出版社

高等学校文科教材

# 情报语言学基础

(增订二版)

张琪玉 著

武汉大学出版社

## 图书在版编目(CIP)数据

情报语言学基础/张琪玉著. —增订2版. —武汉: 武汉大学出版社, 1997. 9

高等学校文科教材

ISBN 7-307-02379-2

- I 情…
- II 张…
- III 检索语言
- IV G254.0

**武汉大学出版社出版**

(430072 武昌 珞珈山)

湖北省武汉第二印刷厂印刷

(430100 武汉市蔡甸区正街176号)

新华书店湖北发行所发行

1997年9月第1版 1997年9月第1次印刷

开本: 850×1168 1/32 印张: 11

字数: 281千字 印数: 1—2000

ISBN 7-307-02379-2/G·322 定价: 11.00元

本书如有印装质量问题, 请寄承印厂调换

## 内 容 提 要

本书是高等学校文科教材之一。

本书系统地讲述了情报语言学的一般问题、情报检索语言的基本理论和基本方法、各种类型情报检索语言的原理、编制法、使用法、性能、在方法上互相吸取的情况、文献分析与标引、自然语言在情报中的应用，以及情报检索计算机化与情报语言学的发展等。

本书可作为高等学校信息管理系图书馆学专业和情报学专业的专业教材，也可供各类型图书馆、情报机构、资料室、档案部门的工作人员学习参考。

## 本版前言

本书1983年由武汉大学出版社正式出版以后,曾被许多高等学校(包括中央电大)的图书馆学和情报学专业采用为教材或主要教学参考书,在社会上也得到比较广泛的流传。

本书虽曾印过8次共96 000册,但现已无处可购得。由于一些学校至今仍采用本书为教材,但储备教材已用完,有的学校不得已而将上届学生用过的教材收回供下届学生再用。为此,决定请出版社再版。

本书1983年以《情报检索语言》为书名初版时,内容共8章。1987年增补了三章和一节,并对各章作了文字上的修改,改名为《情报语言学基础》,再次出版。

本版又增补了“分类法主题法一体化检索语言”和“自然语言在情报检索中的应用”两章,并对其他各章作了文字上的订正,使它能反映情报语言学学科在近年来的发展。

张琪玉

1997年1月23日于上海

## 电大版前言

本书实际上是《情报检索语言》(高等学校文科教材, 1983年6月武汉大学出版社出版)一书的增订本。与原书的区别在于: 增加了第一章 绪论——情报语言学一般问题, 第九章 情报检索语言基本方法, 第十一章 情报检索计算机化与情报语言学的发展; 将第二章改名为情报检索语言基本理论, 并增加了第四节 文献主题的构成因素及层次; 对原有各章作了个别文字上的修改和少量补充, 并对参考文献目录作了补充; 将书名改为《情报语言学基础》。

本书之所以改名为《情报语言学基础》, 主要出于以下几点考虑:

(1) 本版增加了上述章节的新内容, 使全书的全面性、系统性、理论性有所增强, 已基本具备作为一门学科(而不是仅限于一个研究对象)基础知识教材的条件;

(2) 将各种类型情报检索语言乃至自然语言在情报检索中的应用问题放在一起进行统一研究的著作早在70年代初或60年代末即已出现(1972年出版的F. W. 兰开斯特《情报检索词汇规范化》一书是典型的一种), 国内外图书情报界十多年来的研究, 使这方面的知识作为一门学科已经逐步成形。但是, 它至今还没有一个学科名称, 而是使用着“情报检索语言”、“检索语言”、“分类法与主题法”、“文献标引(学)”、“情报检索词汇”等名称。为了确立它在学科之林中的地位, 有必要为它命名;

(3) 目前, 将分类检索语言与主题检索语言以及自然语言在

情报检索中的应用问题作统一研究的必要性在国内还未引起普遍重视。显然,不进行统一研究,我们的视野和思路就不会开阔,理论水平就不可能较快地提高,这对解决情报检索中的语言保证问题,对任何一种情报检索语言的发展,对自然语言在情报检索中应用的发展都是不利的。所以,通过命名,明确提出研究方向,对于这一学科领域各个对象的统一研究可以起“催化”作用;

(4) 关于这一学科领域应怎样命名的问题,我经过了多年的思考。这门学科是对各种类型情报检索语言作统一研究而发展起来的,但是若以“情报检索语言学”命名,则失之偏窄,就不可能将日益发展的自然语言检索包括进来,而将情报检索语言与自然语言检索结合起来研究,则是一种必然的趋势。所以,用“情报语言学”命名,将其定义为一门研究情报检索中的语言保证问题的学科,看来更为确当。

我于1983年将“情报语言学”一词试用于《情报语言学导论研究提纲》和《情报语言学文献分类表》。此后又以“谈谈怎样研究情报语言学”为题,对84级和85级研究生讲第一课,就情报语言学的定义、研究对象、学科形成过程、研究目的、研究范围、研究材料、研究课题、研究方法和相关学科等作了阐述。讲课的提纲曾以《关于情报语言学》为题,编入1985年8月刊印的《〈情报检索语言〉学习辅导》一书第26~28页。本书第一章即为该讲课提纲,仅补充了一些解释。

现遇教材增订再版的机会,将其改名为《情报语言学基础》,目的是借以把这个知识领域作为一门学科提出来,希望分类法的研究者、主题法的研究者、自然语言检索的研究者都汇集到这个研究领域中来,共同推动它更快地进步。

张琪玉

1987年3月10日于武汉大学

## 第一版前言

情报检索语言是专门用于各种手工的和计算机化的文献情报存贮检索系统,表达文献主题概念和检索课题概念的人工语言,包括体系分类法、组配分类法、标题法、单元词法、叙词法、关键词法以及代码语言等多种类型。这类语言之对于图书情报工作,如同自然语言之对于人类社会生活一样重要。检索系统效率的高低,在很大程度上取决于所采用的情报检索语言的质量以及对它的使用是否正确。

从图书情报工作现代化的需要看,作为受过高等学校图书馆学或情报学教育的专业人才,除了必须掌握体系分类法这种传统的情报检索语言的知识外,还有必要掌握更多的情报检索语言基本知识,以便在工作中能正确地使用这类语言,以及向读者普及这类语言的知识,乃至进一步有所研究,结合实际需要改进现用的语言和创制新的语言。

由于情报检索语言知识的用途比较广泛,武汉大学图书馆学系从1977级起,为图书馆学专业和科技情报学专业的本科生、函授生、走读生以及部分研究生、进修生普遍开设了“情报检索语言”课程。从1980年上半年以来,已对800多名学生讲授了该课程知识。本书即是为该课程编写的教材,现被列入教育部主持制订的《高等学校文科教材编选规划》出版。

本书以检索效率问题为中心,讲述和讨论了情报检索语言的基本理论、各种类型情报检索语言的原理、编制法、使用法、性能、在方法上互相吸取的情况,以及文献分析与标引的一般问题,



力求使学习者获得有关情报检索语言的比较完整、系统的基本知识，并对各种类型语言所存在问题的解决有所思考。

本书初稿写成于1980年3月，这是第二次修订本。但是，将各种类型情报检索语言知识统一于一门课程中，连贯起来作分析比较式的讲授，这还是一种探索。编者深知，本书作为一种教材还不够完善，虽经两次修改，一定还有很多缺点。现借出版的机会，恳请专家和读者批评指正，使它能得到改进。

本书在编写和修改过程中参考了一些文献（见参考文献目录），在这里我谨向作者致以谢忱！本书在出版过程中曾得到图书馆学专业教材编审组关懿娴、彭斐章、来新夏等同志，不辞辛苦前来参加审稿会议的曾世荣、韩承铎、范世伟、黄万新、史永元同志，提出书面意见的许培基同志，本系孙冰炎、周继良、杨元生同志，武汉大学出版社刘文华同志等的支持、指正和帮助，对此我表示深切的谢意！

张 琪 玉

1982年12月24日于武汉大学

# 目 录

本版前言	1
电大版前言	1
第一版前言	1
第一章 绪论——情报语言学一般问题	1
第二章 情报检索语言基本理论	5
第一节 情报检索语言在情报检索中的作用	5
第二节 对情报检索语言的共同要求	8
第三节 概念逻辑和知识分类是情报检索语言的基础	15
第四节 文献主题的构成因素及层次	25
第五节 情报检索语言的分类	30
第三章 等级体系型分类检索语言	33
第一节 体系分类法的原理	33
一 体系分类法的构成原理	33
二 类目的划分	35
三 类目的排列	39
四 类目名称及涵义	44
五 学科之间相互关系的处理	47
六 某些图书资料分类的特殊性	49
第二节 体系分类表	51
一 分类表的功用和结构体系	51
二 分类表的编制法	53
三 分类表的设计和编制程序	67

四 分类表的使用法 .....	70
五 分类表的管理 .....	71
第三节 体系分类法的性能 .....	73
一 体系分类法的检索效率 .....	73
二 体系分类法的某些问题 .....	74
第四章 分析-综合型分类检索语言 .....	83
第一节 组配分类法的原理 .....	83
一 组配分类法的构成原理 .....	83
二 “面”的划分和细分 .....	87
三 类号的组配和排列 .....	90
四 组配分类法的类型 .....	91
第二节 组配分类表 .....	93
一 组配分类表的结构体系 .....	93
二 组配分类表的编制法 .....	94
三 组配分类表的使用法 .....	96
第三节 组配分类法的性能 .....	97
一 组配分类法与体系分类法的比较 .....	97
二 组配分类法的某些问题 .....	98
第五章 标题词型主题检索语言 .....	101
第一节 标题法的原理 .....	101
一 标题法的构成原理 .....	101
二 标题及其类型 .....	103
三 标题词的范围及选词原则 .....	105
四 标题词的规范 .....	107
五 标题词相互关系的显示 .....	110
第二节 标题表 .....	118
一 标题表的功用和结构体系 .....	118
二 标题表的编制法 .....	118
三 标题表的使用法 .....	120
四 标题表的管理 .....	127
第三节 标题法的性能 .....	129

一	标题法与体系分类法的比较	129
二	标题法的某些问题	131
<b>第六章</b>	<b>单元词型主题检索语言</b>	<b>138</b>
<b>第一节</b>	<b>单元词法的原理</b>	<b>138</b>
一	单元词法的构成原理	138
二	单元词的范围及选词原则	140
三	单元词的规范	141
<b>第二节</b>	<b>单元词表</b>	<b>144</b>
一	单元词表的功用、结构体系和编制法	144
二	单元词表的使用法	145
<b>第三节</b>	<b>单元词法的性能</b>	<b>147</b>
一	单元词法与标题法、组配分类法的比较	147
二	单元词法的某些问题	149
<b>第七章</b>	<b>叙词型主题检索语言</b>	<b>152</b>
<b>第一节</b>	<b>叙词法的原理</b>	<b>152</b>
一	叙词法的构成原理	152
二	叙词的范围及选词原则	156
三	叙词的组配	159
四	叙词的规范	167
五	叙词相互关系的显示	175
<b>第二节</b>	<b>叙词表</b>	<b>187</b>
一	叙词表的功用和结构体系	187
二	主表的编制法	191
三	辅表的编制法	197
四	叙词表的设计和编制程序	207
五	叙词表的使用法	210
六	叙词表的管理	215
<b>第三节</b>	<b>叙词法的性能</b>	<b>216</b>
一	叙词法与单元词法、标题法、组配分类法、体系分类法的比较	216
二	叙词法的某些问题	218

<b>第八章 关键词型主题检索语言</b> .....	230
<b>第一节 关键词法的原理</b> .....	230
一 关键词法的构成原理.....	230
二 关键词法的表现形式.....	231
三 关键词的范围.....	238
四 关键词的规范.....	239
<b>第二节 关键词法的性能</b> .....	240
<b>第九章 分类法主题法一体化检索语言</b> .....	243
<b>第一节 情报检索语言的兼容化与分类法主题法</b> <b>一体化检索语言</b> .....	243
一 情报检索语言兼容化的重要性及其模式.....	243
二 分类法主题法一体化检索语言的类型.....	245
<b>第二节 分类主题词表的结构原理</b> .....	246
<b>第三节 分类主题词表的编制法和用法</b> .....	249
一 分类主题词表的编制法.....	249
二 分类主题词表的用法.....	251
<b>第四节 分类主题词表的性能</b> .....	252
一 分类主题词表的特殊功能.....	252
二 分类主题词表中分类号与主题词对应的准确性.....	253
<b>第十章 情报检索语言基本方法</b> .....	255
<b>第一节 情报检索语言结构与功能的关系</b> .....	255
<b>第二节 情报检索语言为发挥“对文献的情报</b> <b>内容（及某些外表特征）加以标引”</b> <b>这项功能所采用的方法</b> .....	257
<b>第三节 情报检索语言为发挥“对内容相同及</b> <b>相关的情报加以集中或揭示其相关性”</b> <b>这项功能所采用的方法</b> .....	262
<b>第四节 情报检索语言为发挥“对大量情报加</b> <b>以系统化或组织化”这项功能所采用</b> <b>的方法</b> .....	266

第五节	情报检索语言为发挥“便于将标引用语和检索用语进行相符性比较”这项功能所采用的方法·····	267
第十一章	文献分析与标引问题·····	269
第一节	文献的分析与标引是影响情报处理质量的关键性工序·····	269
一	保证分析标引质量的重要性·····	269
二	评价分析标引质量的标准·····	270
第二节	文献分析与标引的一般方法和规则·····	270
一	分析水平与标引深度·····	270
二	分析角度·····	272
三	分析方法·····	272
四	标引方式方法·····	273
五	导致检索效率降低的分析标引误差·····	274
第三节	文献分析与标引的某些问题·····	275
一	适度标引问题·····	275
二	集中标引问题·····	277
三	利用已有标引成果问题·····	278
第十二章	自然语言在情报检索中的应用·····	279
第一节	自然语言的应用是当今情报检索领域的一种重要发展趋势·····	279
第二节	自然语言在情报检索中应用的各种方式·····	280
一	不标引(无标引)方式·····	280
二	自动抽词标引方式·····	280
三	人工标引方式·····	282
第三节	汉语自然语言检索的特殊问题及汉语分词技术·····	283
一	词典分词法·····	283
二	切分标记分词法·····	285
三	其他分词法·····	285

第四节	单汉字检索	287
一	单汉字检索适应汉语的特点	287
二	单汉字检索的优点和缺点	287
三	关于单汉字检索系统配备后控制词表问题	288
第五节	自然语言检索中各种因素对检索效率的影响	288
一	作为检索依据的文本的类型对检索效率的影响	289
二	检索用语的专指度对检索效率的影响	290
三	在句、段、节、篇不同范围内进行组配检索对检索效率的影响	291
四	文本用词的不规范性对检索效率的影响	291
五	不同标引方式对检索效率的影响	292
六	对自然语言进行词表控制的程度对检索效率的影响	293
七	使用自然语言应注意之点	293
第六节	后控制词表	294
一	“后控制”和“后控制词表”两个概念	294
二	配备后控制词表是提高自然语言检索效率的有效措施	295
三	后控制词表的控制机理	296
四	后控制词表的控制程度	296
五	后控制词表编制上的特点	297
六	后控制词表的几种编制方法	298
七	后控制词表的一种模式	298
八	利用后控制词表的文献检索方法	300
第七节	自然语言与情报检索语言的比较	301
第八节	自然语言取代情报检索语言问题	304
一	问题的由来	304
二	国外研究的进展说明自然语言时代距今尚远	304
三	现实中并未出现放弃情报检索语言的迹象	305
四	自然语言检索中的难题	307
五	自然语言或情报检索语言的未来	307
第十三章	情报检索计算机化与情报语言学的发展	309
第一节	情报检索语言的创新和改造	309

第二节	词表、分类表的机编化和机读化·····	314
第三节	文献标引和索引编制的自动化·····	315
第四节	自然语言检索的实现·····	317
第五节	多种语言的结合使用·····	319
第六节	检索方法的进步·····	320
第七节	情报语言学研究课题和研究方法的变化·····	323
第八节	结束语——情报检索计算机化对情报 语言学发展的深刻影响·····	326
参考文献目录·····		327



## 第一章 绪论——情报语言学一般问题

情报语言学是研究情报检索中的语言保证问题的一门学科，是情报检索学的一个分支。

情报语言学的主要研究对象是情报检索语言，同时也研究自然语言在情报检索中的应用问题。

情报检索语言是体系分类法、组配分类法、标题法、单元词法、叙词法、关键词法以及各种代码体系和引证关系追溯法等的统称。体系分类法等都是针对文献情报内容的检索方法，更确切地说，都是表达文献主题概念和检索课题概念的概念标识系统，是情报检索的语言工具。

情报检索语言的产生可以追溯到古代（如我国汉代刘歆创制的《七略》分类体系）。目前大多数承认的体系分类法和标题法的早期典型，即杜威创制的体系分类法和克特创制的标题法，都是1876年问世的。在此后很长时间内，对这两种检索方法类型的研究，一直是分别进行的。后来又出现了组配分类法、单元词法、关键词法、叙词法，等等。随着图书情报事业的发展，需要对这些方法的优缺点，特别是它们的检索效率进行比较研究。结果，就发现这些方法所要解决的基本问题是相同的，它们的基本原理是一致的，它们在方法上也是互相渗透、互相吸取的，总之，发现它们之间有共同的规律，所以，就有必要对它们进行统一研究。这样，就逐渐形成了情报语言学这门知识。

情报检索语言这一概念大约形成于60年代初或50年代末。情报语言学方面较早的系统著作大约出现在70年代初或60年代末。