

分类语言 与主题语言

张 帆 编著

华中师范大学出版社

37.644
08

分类语言与主题语言

张 帆 编著

华中师范大学出版社

1-11563

鄂新登字11号

分类语言与主题语言

张 帆 编著

*

华中师范大学出版社出版发行

(武昌桂子山)

新华书店湖北发行所经销

华中师范大学印刷厂印刷

开本850×1168 1/32 印张 13 字数 338 千字

1991年12月第1版 1991年12月第1次印刷

ISBN 7-5622-0772-0/H·78

印数 1—1200 定价 3.90 元

内 容 提 要

《分类语言与主题语言》是图书馆学、情报学、档案学专业基础课教材之一，也是国内首次将分类语言与主题语言合为一体的公开出版物。本书系统地介绍了分类语言与主题语言的基本原理、基础知识与功能，文献标引的意义、类型与方法，国内、外分类语言与主题语言发展简介以及情报语言分面组配化、文献标引与情报检索自动化、分类主题一体化的发展趋势等。本书资料新颖、内容丰富、广征博引、说理充分、理论联系实际。

本书可作为高等院校图书馆学、情报学、档案学专业本科生教材或教学参考书，亦可供图书、情报、档案资料部门工作人员自学、实际应用与研究时参考。

前 言

分类语言和主题语言是人们为了贮存和检索文献而创制的两种情报语言，二者不但基于同一逻辑原理，而且总体功能与作用基本相同。本世纪60年代以来，分类语言与主题语言相互引入与渗透、合二为一的趋势日臻明朗。为了适应分类主题一体化的发展趋势，探索我国分类语言与主题语言一体化的途径，华中师范大学图书情报学系在1985年进行课程调整时，便将传统的“图书分类学”与“主题法”合为一门专业基础课“文献分类与主题标引”，并编写了与之配套的试用教材“分类语言与主题语言”及有关教学参考资料。6年多的教学实践表明，将分类法与主题法合为一门专业基础课不仅可行，而且是图书情报专业课程改革的重要内容之一。

《分类语言与主题语言》一书是在该系经过多年试用的教材的基础上，对其内容与体系结构进行全面充实与修订之后编写成的。

本书的特点是，将我国图书、情报界广泛使用的两种检索语言——“分类”与“主题”采用对照和比较的方式，对两者的基本原理及相互关系进行了简明的阐述，详细介绍了分类语言与主题语言的基础知识及文献标引技能。

与此同时，本书广泛吸收了当代国内、外图书情报领域的有关研究成果，参考并引用了许多专家、同行近年来的最新著述。全书共分为6部分：第一部分（第1章）是关于情报语言基础知识的概述，包括情报语言的理论与方法基础、情报语言的类型与性质等。第2部分（第2—3章）是关于分类语言基本概念及分类语言实体（体系分类法与组配分类法）体系结构的介绍，包括文献分类基本概念、分类法编制原则以及分类法（表）结构剖析。第

3 部分(第4—5章)包括主题语言基本概念、叙词表结构等内容。第4部分(第6—7章)以《中图法》和《汉语主题词表》为依据,系统介绍了主题分析与文献标引的方法,其中包括文献标引的意义、作用与种类、文献标引总则及各类型文献标引规则。第5部分(第8—9章)是关于国内、外分类语言与主题语言发展情况的介绍,简介了世界上几部比较有影响的分类法(表)与主题词表。第6部分(第10—11章)是关于分类语言与主题语言性能与检索工具的比较,其中包括对分类语言与主题语言发展趋势的展望。

为便于学习与参考,书中每章均列有复习思考题,书后列有主要参考书目,以供图书馆学、情报学、档案学专业的教师、学生以及图书、情报、档案部门的工作人员参考。

将分类语言与主题语言编写成一部专业基础课教材是一种初步的尝试,旨在对如下问题作些探索:

1. 适应分类语言与主题语言理论与发展的需要,符合多种情报语言融为一体的发展趋势。

2. 便于我国图书、情报部门同时进行分类标引与主题标引的工作。

3. 有助于读者从理论到实践对分类语言与主题语言进行系统研究与探讨,便于将二者进行比较研究。

4. 有利于贯彻教学内容少而精的原则,克服分别讲授时,部分内容重复的弊端。

试用教材《分类语言与主题语言》曾经过武汉大学陈光祚教授、周继良教授及周绍萍副教授指正,在此表示感谢!此外,我还要感谢一直关心本书出版并给予过我很多帮助的老师 and 同行们。

本人学识浅薄,加之时间仓促,手头资料有限,书中难免存在错误和不妥之处,敬请专家、同行指正。

编 者

1991年7月

目 录

1 绪论	(1)
1.1 情报语言概说	(1)
1.2 情报语言的理论与方法基础	(5)
1.3 检索标识与情报语言类型	(20)
1.4 情报语言的基本性质	(26)
2 分类语言概述	(35)
2.1 分类基本概念	(35)
2.2 分类法编制原则	(41)
2.3 分类法(表)的结构原理与性能	(45)
3 分类法(表)结构剖析	(57)
3.1 体系分类法类目表的构成	(57)
3.2 类目的设置与组织	(67)
3.3 体系分类法的标记符号	(80)
3.4 说明与注释	(92)
3.5 类目索引	(95)
3.6 组配分类法类表结构	(101)
4 主题语言概述	(110)
4.1 主题语言发展概况	(110)
4.2 标题法结构原理与性能	(118)
4.3 叙词法(主题词法)结构原理与方法	(126)
4.4 标题法与叙词法性能比较	(142)
5 叙词表结构剖析	(148)
5.1 叙词表结构	(148)
5.2 叙词表选词原则及范围	(163)
5.3 《汉语主题词表》简介	(169)
6 主题分析与文献标引	(190)
6.1 主题分析	(190)

6.2	文献标引的意义、作用与种类	(199)
6.3	检索标目的形式与标引深度	(206)
6.4	文献标引程序与同类文献的区分	(216)
7	文献标引规则	(224)
7.1	文献标引工作总则	(227)
7.2	分类标引一般规则	(231)
7.3	各基本大类文献分类标引	(236)
7.4	主题标引一般规则	(258)
7.5	各类文献主题标引	(267)
8	文献分类语言发展概要	(284)
8.1	中国文献分类语言发展概况	(284)
8.2	国外文献分类语言发展概况	(308)
9	国内、外主题词表简介	(322)
9.1	国外主题词表简介	(322)
9.2	中国词表简介	(338)
10	分类语言与主题语言性能及其检索工具比较	(355)
10.1	分类语言与主题语言的联系与区别	(355)
10.2	分类目录的结构与组织	(365)
10.3	主题目录的结构与组织	(374)
10.4	分类目录与主题目录比较	(377)
11	分类语言与主题语言展望	(380)
11.1	分类语言与主题语言编制原理与方法的改进	(380)
11.2	分类语言与主题语言的自动化与标准化趋势	(386)
11.3	分类语言与主题语言的一体化趋向	(392)
	本书主要参考文献	(406)

1 绪 论

1.1 情报语言概说

20世纪以来,随着科学技术的迅猛发展,“爆炸性”增长的文献情报已经成为信息时代社会进步、科技发展所必须的重要资源。为了适应文献数量急剧增长、科研课题日趋专门化的形势,满足人们对文献情报越来越复杂的检索要求,图书情报部门一方面要广开渠道,多途径地搜集文献情报,扩大文献资源贮存量,同时也要加强对文献情报的揭示与组织工作,为文献情报的科学管理与高效率的交流使用创造条件。

目前,从知识内容的角度揭示和组织文献情报的方法主要有两种:一是采用学科分类号表达文献情报的主题内容与情报提问,按照学科分类的逻辑次序组织检索工具与文献;另一种则是借用自然语言的词语描述文献情报内容,按照术语字顺组织检索工具。前者称为分类法,后者称为主题法。从实质讲,分类法与主题法都是人们为了贮存与检索文献情报而创立的人工语言,这种专门用于文献情报检索的特定语言体系,在图书情报领域称为情报检索语言。

情报检索语言也叫做情报语言、标引语言或检索语言,它是人们汇集、组织与检索文献的手段与工具,是“表达一系列概括文献情报内容的概念及其相互关系的概念标识系统”。情报语言的功能主要体现在以下几个方面:

- ① 用简明、概括的形式描述文献情报主题内容及外表特征;
- ② 聚集主题内容相同的文献情报,展示内容相关或相近的

文献；

- ③ 将大量的文献情报系统化与组织化；
- ④ 提供查检文献情报的途径。

在情报检索系统中，情报语言是联系文献情报源与情报用户的桥梁。图 1.1 给出了情报检索系统的示意图。

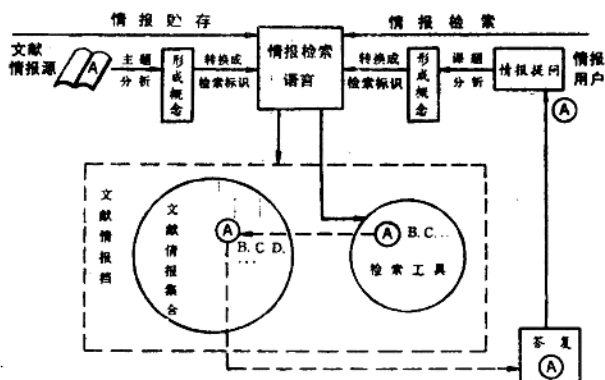


图 1.1 情报检索系统

从上图可以看出，情报检索语言是情报检索系统的核心，是情报贮存与情报检索的语言保障。如果没有作为贮存和检索特定文献而存在的情报语言，各种类型的情报检索要么不可能产生，要么难以获得理想的效果。

1.1.1 情报语言与自然语言

情报语言 and 自然语言一样，有着自己的词汇（字符）和语法。在词汇方面，如用阿拉伯数字“0—9”、字母“A—Z”；或用字母与数字的混合形式；或直接借用自然语言的词语，如“国际贸易”、“超导合金”、“投资”等。情报语言的词汇也叫做检索标识，它是从文献情报主题内容中抽取或概括出来的，是以压缩的

简明形式作为文献情报内容的标签和查索依据而存在的，它是读者查找文献情报的检索入口。

情报语言的词法和句法是指构成检索标识的策略与法规，如复分表使用规则、组配规则、选词规则、标记制度等等。情报语言通过词法和句法来控制与限定检索标识的词型及涵义，使之达到专指、规范与统一的目的。

各种分类表、词表及代码表等是将文献情报内容转换成检索标识的根据，它们是情报语言的“词典”与“词库”。这些“词典”既是各种检索语言的具体体现，也是规定与控制情报语言基本词汇与词型、建立语义网络的法典。目前世界上有2000多种检索语言，如《中国图书馆图书分类法》、《汉语主题词表》、《中国分类主题词表》、《国际十进分类法》、《TEST叙词表》、《LCSH 标题表》等等，它们为各类型图书情报部门揭示文献情报主题内容、建立检索系统、提供查找线索创造了条件。

情报语言为了实施其汇集、贮存与检索特定文献的职能，必须对自然语言中的同义现象、多义现象进行控制与规范，使每个检索标识绝对地与一个具体或抽象的概念唯一单义地相互对应。

和自然语言有所不同，情报语言不强调检索标识的语音，只注重其书面词型；语法简明并条理化。

1.1.2 情报语言、文献标引和情报检索

如前所述，情报语言是情报检索系统的核心，是情报检索工作中特定的交际工具。广义的情报检索包括情报贮存与情报检索两部分内容。简言之，情报贮存就是利用情报语言将文献情报纳入检索系统，使众多文献由分散到集中、由杂乱无序到系统化的过程。在此阶段，一个十分重要的工作环节就是要为纳入检索系统的文献与数据确定检索标识，这种根据检索情报的需要为文献赋予检索标识的工作称为文献标引。文献标引的任务也就是采用恰当的检索标识，将文献情报的主题内容和某些具有检索意义的

外表特征一一记录下来，并组织成查检系统，为用户在大量的文献集合中迅速查检到特定需要的文献情报沟通渠道。不具备检索标识的文献情报不能纳入检索系统。对于特定的文献情报来说，没有输入便没有输出，没有存贮就不可能有效地进行检索。从整体上讲，文献标引是情报检索的基础与前提，是建立检索系统的关键性环节。文献标引的质量和情报检索效率关系极大，准确得当的标引不仅能使相同内容的文献情报有效的聚集在同一检索标目之下，而且便于成千上万的检索标目形成有机的知识网络。反之，不恰当的标引会使相同主题内容的文献分散在不同的检索标目之下，造成系统混乱，影响检索效果。

狭义的情报检索，是指根据读者与用户的特定要求，通过手工或计算机检索工具所提供的线索，从大量的文献情报聚类中查找特定文献的过程。其具体检索方法，可参阅专门书籍，此处不赘述。

1.1.3 情报检索语言概念的形成

早在两千多年以前，我们的祖先就开始采用分类的方法处理文献。此后一千多年，虽然各种类型的图书分类法不断产生与发展，但由于科学文化水平的局限，分类法理论与方法的进展十分缓慢。直到19世纪中叶以后，随着科学技术的巨大进步、科技文献不断增多，图书馆事业才有了迅速发展。在早期的分类法不能满足日趋进步的社会需要的情况下，较为完善的体系分类法——《杜威十进分类法》（简称DC）和主题法的早期形式——美国克特所编制的标题表便于1876年应运而生。此后近百年内，分类语言与主题语言走过了一段分别研究、分头发展的历程。在此期间，分类法和主题法从基础理论到技巧方法均有较大进展，而且相继出现了组配分类法、体系-组配分类法、单元词法、关键词法和叙词法等。到20世纪60年代，随着图书情报事业的迅猛发展，人们检索需要的日趋复杂以及人们认识的拓宽与深化，一个能够统一概括各种分类语言与主题语言及代码系统的理论与方法

的高层次概念——情报检索语言诞生了。由于分类法与主题法基于同一逻辑原理，二者的性质、总体功能及所要解决的基本矛盾相同；现代科学技术的发展对二者的要求相同以及两者在原理与方法上互相渗透、互相影响、互相补充、日趋合一的态势，使情报语言概念的形成及其内涵的日渐丰富成为必需与可能。事实上，情报语言的理论与实践是在逐步适应文献情报需求者准确、迅速、全面地检索所需情报的过程中发展的。当然，分类语言与主题语言也有一些差别，但其差别是在基本原理一致、功能基本相同的基础上产生的。因此，分类语言和主题语言也就有统一研究与学习的意义与基础。

情报语言在情报检索过程中承担着“语言保障”的作用，它是组织与查寻知识的钥匙。特别是进入信息社会以后，情报语言知识的学习与应用已经成为衡量一个社会现代化程度的重要标志之一，也是衡量一个科学工作者学识水平与研究能力的重要依据。对于从事图书情报工作与研究的人员来说，学习并掌握情报传递的特定交际工具——情报语言则是开展工作与获得业绩的前提条件。

1.2 情报语言的理论与方法基础

情报语言是根据情报检索的特定需要而创立的专业语言，其理论与方法建立在哲学、概念逻辑、知识分类与语言学的基础上。对其理论基础进行研讨将有助于加深对情报语言本质特征的认识，有助于掌握其技巧与方法。

1.2.1 情报语言的哲学基础

马克思主义哲学的基本原理之一是存在决定意识，情报语言的产生和发展也遵循这一基本原理。任何一种情报语言都是时代的产物，是社会政治和经济在观念形态上的反应，都反映着一定时期科学技术发展的水平以及人类对文献情报的需求。我国汉代

以儒家思想为指导的《七略》和1876年出现的《杜威分类法》第1版以及克特创制的标题表,虽然都只简单地列举了一些有限的学科与事物主题概念,但这和当时的科学文化水平、人们认识的局限以及检索需要是基本适应的。当前出现的叙词语言以及分类与叙词相结合的检索语言,同样是20世纪人类思想文化与科学技术发展所需要的,是适应信息社会文献数量激增及人们利用文献情报的特点而产生的。认识与理解情报语言的唯物性,有助于人们根据社会发展需要不断改善与优化情报语言性能或创立新型的更加优秀的“语种”,促进科学技术、文化教育发展。

辩证思维是正确认识客观世界的重要前提。各门学科的专家、学者,不管是否采用了马克思主义的术语,实际上都在力图按照客观世界的本来面目,即辩证关系来认识世界,用抽象的思维去概括客体的本质属性。长时期以来,人们对情报语言的研究经过了由个体到总体的描述与概括阶段。即首先对各种具体的情报语言,如体系分类法、组配分类法、标题法、元词法、叙词法等原理、方法和现象进行描述、对其检索效率进行考察后,发现并认识了多种检索语言的共同性质与一般规律。此外,又根据文献情报及课题提问中显示出来的各种偶然性主题与随机现象,探索出一套描述文献情报内容及检索提问的标引规则。这些从实践中归纳总结出来的规律与规则,体现了情报语言在描述文献情报内容、沟通标引人员与检索人员的思想方面所力求达到的内容与形式、本质与现象、整体与局部、组合与分解的辩证统一。

客观世界相互联系的观点是情报语言又一重要思想基础。在整个情报检索系统中,情报语言是作为文献情报主题内容与检索提问的“居间联系”的“中间环节”,即“中介”而存在的。正如列宁在《哲学笔记》中论述事物的普遍联系时所说:“一切都是经过中介(Vermittelt)连成一体,通过中介转化而联系。”^①

^① 李辛生主编,《〈哲学笔记〉简明教程》,浙江人民出版社,1985,285页。

情报语言这个“中介世界”既包含着人类知识的信息，又反映着人类的检索要求；它承担着将文献情报的标引者与检索者的思维差异相互调节与转化、使文献情报的主题内容与检索需求对应一致的职能，其最终目的是促进文献情报中凝固的知识活化成以头脑为载体的智力，成为人们科学劳动的新起点。客观世界相互联系的理论对于情报语言具有两层意义：其一，指各种不同的语种，包括体系分类法、组配分类法、标题法、元词法、叙词法等，在原理与方法上的相互联系、相互依存、相互渗透、相互作用构成情报语言科学的整体；其二，体现在同一种检索语言内，众多的检索标识互相关联、互相制约形成知识网络。此外，相互联系、浑然一体的原理也是文献情报有序化与系统化的基础。为了揭示事物或概念之间相互联系的现象，情报语言或按照学科体系、或按照术语字顺、或采用分类与字顺相结合的方式，使文献情报主题内容系统化与有序化。

此外，正如恩格斯所说：“只要自然科学在思维着，它的发展形式就是假说。”^① 情报语言在自身知识的积累并逐步向着高层理论形态转化的过程中，常常借助于假说与试验的方法，不断提出新的概念范畴，预测发展趋势。例如：分类法与主题法一体化的问题和以自然语言为检索标识的问题等就是如此。

总之，以马克思主义哲学的基本原理来总结、认识和探索情报语言的过去、现在与未来，揭示情报语言一系列现象的基本规律，同时提出相应的新的研究课题，这就为人们从整体上、联系上、相互作用和发展上探讨情报语言提供了重要的理论基础和有效的方法。

1.2.2 情报语言的概念逻辑基础

概念逻辑是一种科学的思维方法，它的任务是将事物之间的本质属性与非本质属性进行区别，并将本质属性抽象成能够采用

^① 《马克思恩格斯全集》第20卷，北京：人民出版社，1971年，583页。

词语表达的概念，以便于交流使用。人们在各个领域内的认识成果，都是通过形成各种概念来加以总结和概括的。各门科学都拥有自己的一系列概念，这些概念是认识成果的标志。例如，几何学就是由“点”、“线”、“面”等一系列概念及其相互关系构成的。

情报语言实质上是一种表达文献情报主题内容的专业符号语言，它在描述文献主题内容和系统展示概念标识及其相互关系方面，普遍地应用了概念逻辑的原理与方法。

(1) 情报语言的概念性。

情报语言的检索标识是以精炼、压缩的形式表现大量知识的手段，也是人们对科学对象与事物的本质属性进行认识、总结与概括的结晶。每个检索标识既可以用来描述文献情报的主题内容，又可以作为文献内容的检索“入口”，或者成为检索课题的一个组配单元。情报语言的概念性要求检索标识概念明确、通用性能好，具有研究与检索的意义。与此同时，检索标识还要避免含义不清、概念重复、矛盾和交叉的现象。从聚集相同文献的特定职能出发，情报语言必须对自然语言的多义性和同字异义等现象进行控制与规范，以克服同一概念用词不一或不同概念用词雷同的情况。因为前者会造成漏检，后者易造成误检。对情报语言标识进行规范与控制的目的，是使检索标识和表达事物的概念相互对应。这种建立在事物、概念、词语三者严格一致基础上的关系，集中体现了情报语言的概念性。

(2) 依靠概念之间的外延关系建立知识网络。

构成情报语言的检索标识如同表达大千世界万事万物的各种概念一样，每个标识不仅具有一定的词语含义，而且相互之间存在着各种各样的语义关系。根据外延是否部分重合的标准，可以将概念标识区分为相容关系和不相容关系两大类型。如果两个概念的外延至少有一部分相互重合，这两个概念即互称为相容概念。相容概念又可分为等同关系、属种关系和交叉关系三

种类型。如果两个概念的外延完全不重合即称为不相容关系。不相容关系又可分为矛盾关系、并列关系和对立关系。概念之间的并列关系既有相互兼容的形式，也有互不相容的表现。

① 等同关系。

等同关系又叫做同一关系、全同关系或用代关系，即指两个概念标识的外延全部重合的关系。等同关系的概念标识所表达的内涵与外延完全相同或相近，以致于可以互相取代。比如：“蕃茄”与“西红柿”，“符号逻辑”与“数理逻辑”，

“放射生物学”与“辐射生物学”，“陆游”与“陆放翁”等均为等同关系。等同关系一般体现为同义词、学名与俗称等概念关系。具有等同关系的概念互称为等同概念。如图 1.2 所示，M 圈与 N

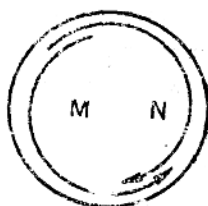


图 1.2 等同关系

圈重叠，表示两个外延完全等同。

② 属种关系。

属种关系又称为等级关系、属分关系，是指一个概念的外延包含着另一个概念的全部外延的关系，也是广义概念同狭义概念之间最邻近的隶属关系。比如：“生物学”与“植物学”，“动物”与“脊椎动物”，“宇宙”与“地球”等等均为属种关系。

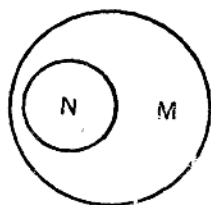


图 1.3 属种关系

其中“植物学”、“脊椎动物”及“地球”等概念的外延分别包含在“生物学”、“动物”与“宇宙”等概念的外延之中。外延较大的概念称为属概念或上位概念；外延较小的概念称为种概念或下位概念。情报语言检索标识之间的上位概念与下位概念、整体概念与部分概念均表现为概念的属分关系。如图