

汉语主题词表

标引手册

2-62



科学技术文献出版社

G 354.22 - 62

Q

43_{c.1}

《汉语主题词表》标引手册

钱起霖		主 编
王剑雄	杨霞丽	郑亚晖
赵宗仁	高崇谦	编 著
朱孟杰		审 定

科学技术文献出版社

1985

213124

内 容 简 介

本书详细介绍了《汉语主题词表》的结构和使用方法，重点阐述了文献主题标引的方法和步骤，并以实例说明如何进行主题分析、如何选用标引词以及标引应遵循的各种规则。书中对于编制主题目录和主题索引的一般方法，特别是对建立计算机化文献检索系统的文献标引问题及其具体形式进行了探讨。书后附有标引实例、名词术语和有关的参考文献。

本书是配合《汉语主题词表》的应用而编写的实用手册，是情报、图书、档案管理以及检索刊物编辑部必备的一种标引工具书，也可作为高等院校图书情报系教学实习的参考书。

《汉语主题词表》标引手册

钱起霖 主编

科学技术文献出版社出版

中国科学技术情报研究所印刷厂印刷

新华书店北京发行所发行 各地新华书店经营

*

开本：787×1092¹/₃₂ 印张：9.5 字数：200千字

1985年2月北京第一版第一次印刷

印数：1—20065册

科技新书目：91—56

统一书号：17176·419 定价：2.15元

前 言

当前，我国的科技情报工作正在进入一个新的发展阶段。为了适应国家四化建设需要，科技情报部门与图书馆开始利用电子计算机对大量文献进行存贮与检索，着手研制汉字信息处理系统及建立中文文献的计算机化检索系统，对于手工检索系统也要求有新的发展和提高，即必须相应地由传统的单一分类体系，向着分类与主题体系并行且互为辅助的方向发展。这一发展趋势决定了必须开展主题法的应用和推广工作。

一九八〇年底，我国情报界与图书馆界联合编制的一部大型综合性《汉语主题词表》问世。近年来许多专业性主题词表，如：《国防科学技术主题词典》、《电子技术汉语主题表》、《航空科技资料主题表》、《原子能科技资料主题词典》、《机械工程主题词表》、《农业主题词表》、《铁路汉语主题词表》以及《常规武器专业主题词表》等陆续出现，为主题法的应用奠定了基础。尽管如此，对广大情报工作人员与图书管理人员来说，使用主题法，特别是使用主题词法(叙词法)来加工处理文献，还是一项比较陌生的工作。有了主题词表，如何利用词表来标识文献、标引中应遵循那些规则、如何编制检索工具等等，是主题法应用阶段中迫切需要解决的课题。本书正是为解决这些课题而作的有益尝试。

《汉语主题词表》是实现情报检索必备的工具，也是国家推荐在全国统一使用的工具书。为推动该书的应用和推广工作，我们编写了这本文献标引工作手册，作为该书的使用指

南。这本手册共分十章，内容包括词表的结构和使用方法，标引工作方法和规则以及检索工具编制方法等。书中除对这些问题作了比较全面系统地阐述外，还以实例作了示范说明，拟从理论和实践两方面对读者进行指导。

本书不仅适用于手工检索系统文献标引工作，而且对计算机检索系统的文献标引工作也具有现实意义。本手册的出版，无疑将为情报、图书、档案部门组织文献主题标引、编制主题目录和索引，建立计算机文档以及开展情报检索服务开辟了新的途径，并为进一步完善手工检索系统，建立机器检索系统，统一全国文献检索体系创造了条件。同时，也将为培养一支文献标引队伍提供一部实用教材。

本手册在编写过程中，曾参阅了大量的国内外科技文献，并使用《汉语主题词表》作了实际的标引试验，从而为本手册的编写提供了理论基础和可靠的数据。本手册的初稿，曾在北京、上海、四川、新疆等地和冶金、煤炭、农林、中国科学院等部门作过试讲，并得到一些应用部门和专家们的指正，在此谨致谢意。

由于编者水平有限，书中必定存在不少问题，乃至错误，敬请读者批评指正。

编者

一九八四年二月

目 录

第一章 概论	(1)
第一节 情报检索与标引工作.....	(1)
第二节 标引工作与检索语言.....	(4)
第三节 主题标引的规范化.....	(10)
第四节 标引人员的素质.....	(12)
第二章 汉语主题词表的体系结构	(15)
第一节 词表的体系结构.....	(16)
第二节 主题词定义及其选择原则.....	(18)
第三节 主题词的参照系统与语义网.....	(24)
第四节 主题词款目结构与主题词表的组成.....	(37)
第三章 词表的查找方法	(56)
第一节 主表与辅助索引之间的联系.....	(57)
第二节 查表步骤与方法.....	(59)
第三节 词义的确定与扩充查词.....	(64)
第四章 主题标引的步骤与方法	(69)
第一节 主题分析概述.....	(69)
第二节 主题的类型与结构.....	(71)
第三节 文献主题分析提纲.....	(76)
第四节 主题分析步骤.....	(77)
第五节 主题分析误差与注意事项.....	(82)
第六节 选定主题标引方式.....	(87)
第七节 主题概念的转换.....	(92)
第八节 标引记录.....	(97)

第九节	标引结果的审核	(106)
第五章	标引工作总则	(110)
第一节	标引工作一般规则	(110)
第二节	标引词的书写规则	(112)
第三节	标引的专指性规则	(113)
第四节	通用概念词的使用规则	(116)
第五节	直接上位词和靠词标引规则	(117)
第六节	标引深度规则	(118)
第七节	过度标引的控制规则	(120)
第八节	标引的一致性规则	(121)
第九节	新词的推荐规则	(122)
第十节	标识词使用规则	(123)
第六章	主题标引的组配规则	(125)
第一节	组配原理	(125)
第二节	组配形式	(128)
第三节	组配规则	(132)
第七章	特殊专业的组配标引规则	(142)
第一节	化合物的组配标引细则	(142)
第二节	金属材料组配标引细则	(156)
第八章	标题与标题词的选定	(160)
第一节	标题的定义与作用	(160)
第二节	标题的结构类型	(160)
第三节	标题的级别	(163)
第四节	标题词的选定原则	(165)
第五节	标题的组配符号	(176)
第九章	主题目录与主题索引的编制方法	(178)
第一节	选定引得深度	(179)

第二节	组织安排标题体系·····	(181)
第三节	主题目录的编制方法·····	(187)
第四节	主题索引的编制方法·····	(194)
第十章	机检文献档的主题标引 ·····	(200)
第一节	标引形式与著录工作单·····	(200)
第二节	标引深度与自由词的使用·····	(206)
第三节	标引词的分组与连符的使用·····	(208)
第四节	机编主题索引的标题符号·····	(211)
第五节	上位词登录·····	(215)
附录：	·····	(220)
I	标引实例·····	(220)
I	名词术语注解·····	(277)
II	参考文献·····	(294)

第一章 概 论

随着现代科学技术的飞速发展，科技文献的数量急剧增长，种类繁多，内容庞杂，而且代谢频繁。因此，文献中的某一项情报常常可能是许多专业课题的共同需要，而一个专业或课题又往往需要从多种文献中获得所需的情报。这就向情报检索工作提出了新的要求。为了解决大量的复杂的情报来源与情报用户的特定需求之间的矛盾，近三十多年来国内外的情报工作者对于情报检索技术进行了大量的研究和实验。特别是电子计算机和其它一些新技术在情报检索工作中的实际应用，推动了情报服务工作的的发展，使情报检索形成情报学中的一个主要学科分支。

进行情报检索，需要有情报检索系统。但无论是组织手工检索系统或计算机检索系统，首先要解决的问题是需要对大量的文献资料进行加工处理，只有把大量文献按一定的体系和规则加以处理之后，才能将它们有效地组织成检索系统。但是，目前这一环节恰恰是相当薄弱的，尤其是文献的主题标引工作，在我国还刚刚开始，远远不能适应新形势发展的需要。为此，开展文献主题标引，推广用主题法标引文献，乃是情报图书部门一项最紧迫的任务。

第一节 情报检索与标引工作

情报检索是开展情报研究与情报服务的基础，其目的是

为了解决大量复杂的情报来源与情报用户的特定需求之间的矛盾。从广义上讲，情报检索是指情报的加工、贮存以及针对用户的特定需求查找出所需情报的全过程。

情报检索过程包括两个基本环节：一是要把搜集的情报源，即文献、数据等按一定的体系和规则进行分编处理和贮存，其中包括**标引**工作和编制**检索工具**的工作，从而构成**检索系统**；一是针对情报用户的需求，利用手工或机械的方式由检索系统的检索工具中查出所需情报的线索提供给用户，或依照线索由文献库中提出原始文献，从而形成**检索输出**。

进行情报检索，要依靠情报检索系统。但无论是组织手工检索系统或计算机检索系统，首先要解决的问题是需要对大量的文献资料或数据进行前期的加工处理。只有把大量文献按一定的体系和规则加以处理之后，才能形成有效的检索系统。

文献的前处理工作，主要是为纳入检索系统的每一篇文章**文献(或每一项数据)**提出**检索标识**。检索标识，是以简练的形式表示的文献特征，目的是用以区分和辨识文献，作为有序贮存和检索文献的依据。无检索标识的文献，不能形成检索系统，也不能有效地对之进行检索。

文献的特征分为**外表特征**和**内容特征**。文献的篇名(题目)、作者姓名、出处、出版者、报告号、专利号等，属于外表特征。文献中所论述的问题、观点和见解等，属于内容特征。以文献的外表特征直接作为检索标识者，称为**自然标识**，这种标识不能提供有关文献主题内容的确切线索，只提供某种特定的检索途径。以文献的内容特征作为检索标识者，称为**人为标识**。人为标识能够提供有关文献主题内容的线索，并且所提供的检索途径恰恰是最大多数情报用户普遍

需要的一种形式。提出人为标识时，须要对文献的内容进行主题分析，并且要加以高度概括和压缩。

情报检索需要借助于**检索工具**。检索工具是指以压缩形式存贮和提供查找文献资料线索的一种工具。其中主要包括三部分内容：作为检索途径的检索标识、文献的题录（包括简介或文摘）以及文献的地址（馆藏号码、报告号或专利号等）。**检索工具**是检索系统的具体体现。在检索工具中，文献资料线索按检索标识的异同分别编排，检索标识之间按一定的体系顺序加以排列，从而构成**检索款目**。

标引工作就是给文献资料赋予检索标识的处理过程。经过标引，文献资料获得检索标识，由无序集转化为有序集，从而为编制检索工具、进行文献检索提供了依据。因此，可以说标引工作是情报检索的关键环节。

正确的标引不仅使同一主题的文献线索恰当地集中在一条检索款目之内，而且使检索款目之间形成有机的整体，便于检索。反之，标引不当会造成误差或错误，将使属于同一主题的文献线索分散到许多检索款目之下，造成检索混乱，甚至会使整个检索系统失去应有的效用。可见，情报检索工作的质量，主要取决于标引工作质量的优劣。无论组织手工检索系统或计算机检索系统，情况都是如此。

标引工作可以把文献作者、标引人员与情报用户三者联系起来。因为标引工作既能表达出作者意图，又能使之与用户的需求相对应。从实质上讲，标引工作是为情报用户能在大量文献中迅速查到特定需要情报建立沟通渠道。

在组织建立情报检索系统的过程中，特别要重视标引工作，研究标引理论和技术，只有首先组织好文献标引，才能为情报检索服务打下坚实的基础。

第二节 标引工作与检索语言

为了便于检索，文献资料的具体检索标识必须具有较小的容量，更重要的是还必须具有准确的语义概念，即应以压缩的形式来表达文献的内容特征。因而文献的检索标识需要用语言形式来表示，这种特殊的语言就称为**情报检索语言**，简称**检索语言**。

近升年来，在情报检索这个分支学科中，对检索语言的研究有了很大的进展，开始引起很多情报学和图书馆学研究人员的注意。特别是由于五十年代末主题词型(即叙词型)检索语言的出现和发展，使得情报检索计算机化不但成为可能，而且实际上已经取得了很大的成果。因此，研究和了解检索语言的理论和实践，就成为情报工作人员，特别是从事标引工作人员的一项基本任务。

1. 检索语言与标引形式

检索语言是一种专门的人工语言，在标引工作中用来描述文献的内容特征(或外表特征)，从而形成检索标识；在检索过程中用来描述检索提问，从而形成提问标识。检索语言亦称标引语言。

情报检索语言既然是一种人工语言，那么就 and 人们日常交往所使用的自然语言有所差别。人们在处理文献资料时，首先遇到的是自然语言。要把用自然语言书写的文献内容用检索语言来概括描述，就需要有一种类似词典那样的工具，称为**检索词典**。检索词典中具体体现了检索语言所用的**字符、词汇及表达复杂概念的语法**，包括构词法和句法。

情报检索语言有许多种类型，但任何一种检索语言对于文献来说都是一种**分编体系**。因为，用某一种检索语言为某一篇文献给出检索标识之时，就意味着把这篇文献**归入**具有同一检索标识的一个子集合(或类目)之内。

检索语言基本上分为两大类型，即**分类语言**和**主题语言**。分类语言和主题语言都是在实践中产生和发展起来的，两者既有共同点，又有区别，既各有侧重，又相辅相成。

所谓**分类语言**，是指以数字或以拉丁字母和数字混合作为基本字符，采用字符直接链接并以圆点作为分隔符的书写法，以基本类目作为基本词汇，以类目的从属关系来表达复杂概念的一类检索语言。其检索词典即具体的分类表。例如《中国图书资料分类法》，是分类语言的一种体现。它以拉丁字母A-Z和数字0—9作为字符，用“a”、“-”、“/”、“{ }”、“+”、“·”、“()”、“=”、“< >”、“《 》”等通用复分号加以辅助。用字符链接、每三位数字后(不计开头的拉丁字母)加一圆点为构词法，用类目级别的从属关系表达复杂概念。

那么如何使用这种语言对文献进行标识？现举几个实例加以说明。

例1：《**可锻铸铁冶炼**》一文，通过查阅《中国图书资料分类法》类表的“冶金工业”类内，可以查到“TF593.1”这个类目，语义概念是“可锻铸铁”。于是即可用“TF593.1”来作为本文的检索标识。TF593.1除表示这一具体事物概念外，还含有冶炼的意义，这一层含义是由类目级别的从属关系中所决定的。如：

TF593.1	可锻铸铁
TF593	铸造生铁
TF59	炼铁产品

TF 5 炼铁

TF 冶金工业

T 工业技术

上例中，我们通过向上级类目逆查时，可以看到在“TF-59”这个类目之下，有一条注释，其中明确指出：“专论某一种铁的冶炼方法入以下各类”。这就更加明确的指出，用“TF593.1”这个语义符号标识《可锻铸铁冶炼》这一篇文章是十分确切的。

例2：《中国水稻育种经验》一书，使用《中国图书资料分类法》标识时，同样在类表中可以查到“S511.1”是水稻的含义。而“水稻育种”在类表中没有现成的标识可用。但在“S5农作物”下有一专用复分表，其中复分号“03选种、育种”可用。于是我们可标识为“S511.103”。然而尚有“中国”这个概念须要表达，根据类表中规定的国家区分符号“()”加上“中国”地区号，就成为“S511.103(2)”。

这个标识，在构成过程中应用了检索语言的语法。而这种语法是该检索语言中规定的。同样从级别从属关系中可以看出它的准确性。

例：

S511.103(2) 中国水稻育种经验

S511.1 水稻

S511 稻

S51 禾谷类作物

S5 农作物

S 农业科学

从以上两例中可以看出，分类语言中概念是由人为的符号表示的，而这些符号都含有语义。在上下级从属关系中，

上位类的语义限定了下位类的语义。

这种检索语言的特点是以科学分类为基础，比较能集中体现学科的系统性，反映事物之间的学科从属关系，便于按科学门类进行**族性检索**。它的基本体系结构是按知识门类的逻辑序列，从总到分，从一般到特殊，从低级到高级，从简单到复杂，逐级展开的层累号码制。给予文献的检索标识称为**分类号**。

所谓**主题语言**，是指以自然语言的字符为字符，以规范化或未经规范化的名词术语为基本词汇，以概念之间的形式逻辑作为语法和构词法，用一组词语作为文献检索标识的一类检索语言。

主题语言又分为标题语言、元词语言和主题词语言(又称叙词语言)等。主题语言的检索词典是具体的词表，根据类型分别称为标题表、元词表、主题词表(又称叙词表)等。

《汉语主题词表》所体现的检索语言是主题词语言。它是用规范化的汉语词汇经过优选之后作为主题词，每个词汇在词表中规定为单义词，以词间的**形式逻辑**关系作为语法和构词法。通过以下实例，可以看出使用此种检索语言对文献进行标识的形式。

例3：《**利用电阻率异常变化预测地震的可能性**》一文，用以下一组主题词标识：

地震预报

地电异常

相关分析

地震前兆

电阻率

大地震

这一组主题词，是通过查找《汉语主题词表》之后，根据一定的规则选出作为本文检索标识的。“利用地电阻率的变化异常现象来预测地震”的这一主题，可由“地电异常”、“电阻率”、“地震预报”、“相关分析”、“大地震”、“地震前兆”几个主题词确切加以表达。主题词之间的逻辑关系与词的排列顺序无关，而是以词间的形式逻辑关系为基础。

例4：《利用生物钟提高抗癌药的效力》一文，其检索标识为：

抗癌药
药效
生物钟

从上例中可以看出，主题概念是用规范化的词语来加以表达的。这是主题词法的特点。

《汉语主题词表》的体系特点是以词的自然序列(音顺)排列，着眼于主题词所表示的特定概念，反映与之有关的部分或全部问题。这种具体的检索语言，以主题词作为文献标识单元，以词间形式逻辑关系表达语义。此种检索标识比分类号更直观、易记，易于表达复杂的概念，具备灵活性和专指性。如例4一文，用《中图法》标识时，可用“R979.1”标识，语义是“抗肿瘤药物”，然而抗癌药效力与生物钟之间的相关关系就很难表示出来。用主题词标识则概念明确，并揭示出主题内容而满足检索要求。

使用不同的检索语言标引文献，所得的检索标识形式不同，效用也有差异。

2. 主题标引与主题词表

标引工作按使用检索语言的类型分为**分类标引与主题标**

引两种。主题标引是指以词或短语作为文献检索标识的一种标引方法。

主题标引按使用词的规则又可分为**自由标引**与**受控标引**两种类型。

自由标引又称抽词标引，是指不借助检索语言词典而直接将文献中关键性的词或短语抽出来作为检索标识的一种标引方法。由于自然语言中的词汇存在着大量的同义词、多义词，易于造成用词不统一或概念含混，从而会使属于同一主题的文献分散，因而检索系统会出现混乱。这种标引方式，对于包含大量文献的检索系统，还未见有采用的。特别是对于计算机检索系统除早期为编制题内关键词轮排索引外，目前的各种实用文献数据库已很少使用此种标引方式。

受控标引又称赋词标引，是指采用统一规范化的检索语言的一种标引方法。其特点是借助统一的检索词典，对用词的词义加以限定。检索词典即词表，具有严谨的体系结构，并对同义词、多义词等加以控制。标引时选用词表中收录的主题词作为文献的检索标识。

受控标引的优点是能在统一用词的基础上使同一主题的文献相对集中，从而使检索系统有序化和规律化。但标引时则必须把文献中表达主题的自然语言**转换成**规范化的检索语言。这种转换类似于翻译，需要进行语义分析和概念对比。如上述例3中的“**地电异常**”与“**相关分析**”两个词，就是对文献内容进行语义分析再与词表中词的概念对比后选定的两个词。

目前国内外大多数检索系统都是利用主题词表进行受控标引的。主题词型情报检索语言，是五十年代出现的一种**新型检索语言**。它在发展了传统标题语言与元词语言的基础