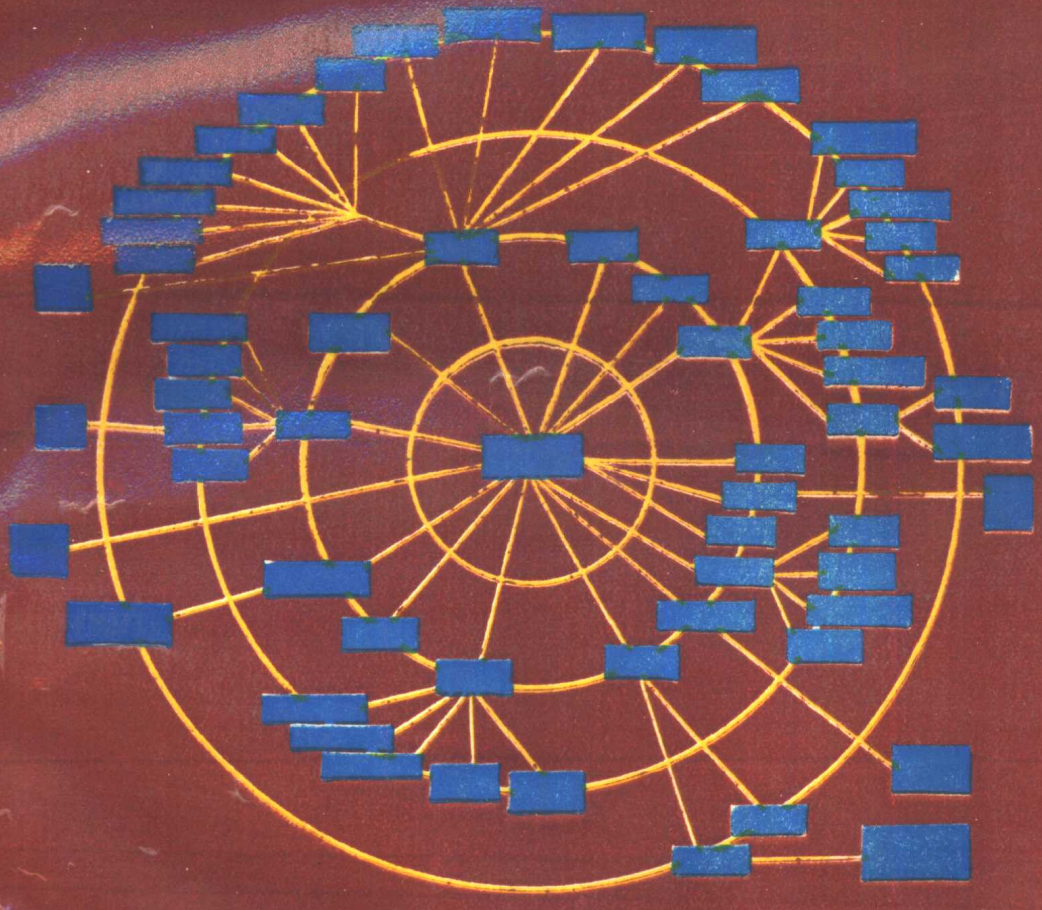
[美]F·W·Lancaster 著

情报检索词汇控制

侯汉清　戴维民　陆宝树　译



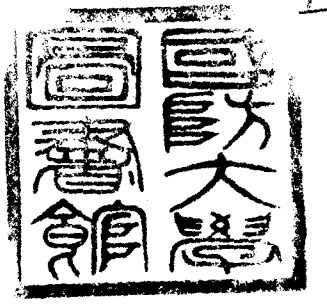
同济大学出版社

*0584735
2 032 9153 4

情报检索语言丛书

情报检索词汇控制

〔美〕 F·W·兰开斯特 著
侯汉清 戴维民 陆宝树 译
陆宝树 校



同济大学出版社

（沪）204号

内 容 提 要

本书对情报检索中词汇控制的各方面问题作了比较全面的论述，尤其是对叙词表的原理和方法作了更为详细的阐述。同时对情报检索词汇控制的新方法、新问题，如分面叙词表、词表的兼容与互换、自然语言检索、计算机在词汇控制中的应用和混合系统等也作了评述和分析。

本书根据原著第二版译出，共二十二章。另有作者简介、译者前言和第一版梗概。本书可供图书馆学、情报学和档案学专业的师生阅读，也可供广大实际工作者和科研人员使用。

责任编辑 张智中
封面设计 李志云

Vocabulary Control for Information Retrieval,
Second Edition
Arlington, Virgiria, Information Resources Pr., 1986
F·W· Lancaster
情报检索语言丛书
情报检索词汇控制
[美]F·W·兰开斯特 著
侯汉清 戴维民 陆宝树 译
陆宝树 校
同济大学出版社出版
（上海市四平路1239号）
同济大学印刷厂印刷
新华书店上海发行所发行
787×1092 1/16印张 12.25字数 310千字
1992年8月第1版 1992年8月第1次印刷
印数 1—3500 定价：7.10元
ISBN7-5608-0977-4/Z·44

作 者 简 介

弗雷德里克·威尔弗雷德·兰开斯特(Frederick Wilfrid Lancaster) 是一位对美国图书馆学、情报学发展有一定影响的学者,又是一位著名的图书馆学、情报学教育家。

兰开斯特于 1933 年 9 月 4 日出生在英国达勒姆城。1950—1954 年在美国泰因河畔纽卡斯尔图书馆学院学习。毕业后先后任纽卡斯尔公共图书馆高级助理、阿克伦公共图书馆高级馆员、华盛顿荷纳公司系统评价组组长、国家医学图书馆副馆长特别助理及韦斯泰特研究公司情报检索服务部主任等职务。1970年后在伊利诺斯大学图书馆学、情报学研究生院执教、主讲情报存贮与检索、情报系统与情报机构评介、主题分析和文献计量学等课程。1972年由高级讲师升任教授。他是美国情报学会分类法研究专业组的成员,还是伦敦分类法研究小组成员。

兰开斯特通过写作、教学及答复咨询为图书馆学及情报学的发展作出了重要的贡献。他勤于写作,著述甚丰,除了发表了近百篇专业论文及报告外,还撰写了下列主要专著:

①《情报检索系统:特性、试验与评价》(*Information retrieval systems; characteristics, testing and evaluation*, 1968),1978 年出第二版,先后译成日文和俄文,荣获美国情报学会1970年颁发的最佳图书奖。

②《医学文献分析检索系统咨询检索服务的评价》(*Evaluation of the MEDLARS demand search services*, 1968)。

③《情报检索词汇控制》(*Vocabulary control for information retrieval*, 1972),1986年出第二版,先后译成日文、俄文、阿拉伯文和西班牙文。

④《联机情报检索》(*Information retrieval on-line*, 1973)荣获美国情报学会 1975 年颁发的最佳图书奖。

⑤《图书馆服务的测试与评价》(*The measurement and evaluation of library services*, 1977),荣获美国图书馆协会1978年颁发的"拉尔夫·肯奖"。

⑥《走向无纸情报系统》(*Toward a paperless information systems*, 1978)。

⑦《情报系统与情报机构的兼容问题》(*Compatibility issues affecting information systems and services*, 1983)。

⑧《电子时代的图书馆和图书馆员》(*Libraries and librarian in an age of electronics*, 1982)

⑨《叙词表的编制和使用简明教程》(*Thesaurus construction and use; a condensed ensed course*, 1985)。

⑩《如果你想评价你们的图书馆……》(*If you want to evaluate your libraries…*, 1988)。

此外他还曾经为联合国教科文组织、粮农组织、美国中央情报局、应用语言学中心、美国国家医学图书馆及澳大利亚国家图书馆进行过咨询服务,为他们编写过情报系统与情报机构的评价指南或评价报告。其中《MEDLARS 工作效率评价报告》荣获美国图书馆协会1969年颁发的最佳文献工作论文奖。

兰开斯特的著作主要研究情报检索系统的智能和概念结构方面的基本问题。他的主要建树是在情报检索语言、情报系统的评价及系统和用户的交互作用等方面。在他的研究成果中影响最大的是他在英国克兰菲尔德试验的基础上对情报系统性能评价标准和方法的发展。他的《图书馆服务的测试与评价》一书填补了图书馆学研究中的空白。他最有争议的著作是《走向无纸情报系统》，该书发表后在图书馆及情报界引起了热烈的争论。兰开斯特治学严谨，善于把严密、周全的方法与清晰的表达结合起来，深入浅出地阐明一些较复杂的情报检索的概念，使之宜于被图书馆与情报专业的学生及实践工作者所理解。这一特点使其著作具有更加广泛的影响。

兰开斯特曾多次来华参观和讲学，他的著作在中国享有较高的声誉，有着广泛的影响。上述著作中①、③、⑥、⑦、⑧、⑨六种已经译成中文出版，受到中国图书馆及情报界的欢迎，不少已被选为大学的教材及教学参考书，或被用作图书情报工作的指南。

此为试读，需要完整PDF请访问：www.ertongbook.com

译 者 前 言

(一)

《情报检索词汇控制》是美国著名情报学家 F·W·兰开斯特(详见作者简介)的代表作之一，第一版于1972年出版。本书全面论述了情报检索词汇控制的各方面问题和各种方法，在图书馆及情报界有着较大的影响，先后被译成日文、俄文、阿拉伯文、西班牙文和中文，被认为是情报检索领域的重要著作。

从70年代到80年代的10多年中，情报技术的发展日新月异，电子计算机的一代代的更新，联机数据库的蓬勃兴起，检索技术和模型的日趋完善，使得情报检索过程中的词汇控制方法也在不断发展，情报检索语言的一些基本方法中渗入了新的手段和技术。正是在这种背景下，F·W·兰开斯特全面总结了情报检索词汇控制的新成果，并精心提炼后熔进了他的著作，这便是1986年出版的《情报检索词汇控制》第二版。

与其说本书是《情报检索词汇控制》的再版，还不如说是一部新的关于词汇控制的著作。尽管本书的两个版本有许多继承性，但是该书的核心内容和写作方法都有了许多根本性的变化。这种变化的实质就是情报检索词汇控制方法的新进展。

(二)

让我们来简略地了解一下本书的基本体系和主要内容。

全书共分22章，比较全面地阐述了情报检索词汇控制的目的、作用、方法以及新趋势和新进展。

情报检索系统由标引和检索两大部分组成。在两大部分中要分别将文献和用户提问的概念分析，转换成系统语言，这个过程需要对词汇进行必要的控制，以达到有效地检索出所需文献的最终目的，这也就是词汇控制的目的。词汇控制方法既适应于先组系统，也适应于后组系统。现代情报检索系统的主流则是后组式的，因此，本书主要论述了后组系统的词汇控制(第1、2章)。

情报检索词汇控制，以叙词表而言，它首先包括下列步骤：(1) 从文献和用户提问中收集词汇；(2) 对词汇进行有效的组织，包括词汇之间等级关系和相关关系的建立；(3) 词形控制和复合词的处理；(4) 同义词处理，建立入口词表；(5) 同形异义词控制，即对词义进行控制，采用范围注释等方法。通过这些过程，那些自由的、游离态的词得到了控制，我们得到的是规范化的标引词或检索词。将这些词组成一个有序系统，便是受控词表。

一般来说，一部较好的受控词表应包括字顺显示和系统显示两个互为补充的组成部分。字顺显示在传统的词表中是一种主要的显示方法，在现今相当一部分词表中仍然是一个主要部分。但系统显示在词表结构中的地位已日益提高。与此同时，人们试图将字顺显示和系统显示组成为一个有机结构，英国学者研制的《分面叙词表》就是其范例。F.W. 兰开斯特赞许地指出："在某种意义上说，《分面叙词表》堪称最好的叙词表。这受益于周到的分面分析，因而

能协调地显示最重要的词间关系和协调一致地控制同义词"(第11章)。因此，分面分类法引入了叙词表，不仅大大改进了词表的系统显示方法，而且更重要的是改变了整个词表的结构。在词表显示中，范畴显示，词族(等级)显示，词形显示等都是有效的辅助显示方法，也是词表常用的显示方法(第3—4，6—11章)。

叙词表的发展，导致了词表编制标准和准则的诞生。这些标准和准则又有效地控制着数以千计的词表的结构。国际性的单语种叙词表编制标准(ISO2788)和多语种叙词表编制标准(ISO5964)，在促进世界范围内的词表规范性和兼容性方面起了积极的作用(第5章)。

词表是一个相对稳定的系统，但文献和主题则是一个动态系统。要使词表不断地适应标引和检索的需要，应定期进行更新和不断扩充，这也是一部词表保持其实用性的基本要求(第12章)。

词汇控制既是一个智力过程，同时也是一个事务性操作过程，需要耗费大量的人力和物力。当计算机介入之后，则使词汇控制变得便捷多了，编表人员可以从大量繁杂的编表事务中解脱出来。在叙词表编制过程，直至叙词表的使用(标引和检索)中，都可以充分利用计算机这一有效的工具。这导致了现代词表机读化的趋向(第13，21章)。

词汇控制也是一个系统工程，因此既要保证词表的编制质量（宏观结构和微观结构的优化），同时也要充分考虑词汇控制的成本效益，还要使词表在提高检全率和检准率方面起到积极的作用。当然，一部理想的词表应该做到成本低，质量高，检索效率理想。然而，这几者往往又是不可协调的。在实际过程中，通常要有所舍弃。当然，词表的使用方法、标引员和检索者的技能，也是影响因素之一(第15，16，22章)。

尽管叙词表是词汇控制的主要方法之一，但是并不意味着所有情报检索系统都采用这一方法。实际上，现今不少数据库没有完全采用，甚至不采用叙词表方法，而采用了其他的词汇控制方法，例如自然语言检索采用后控词表，或将受控语言和自然语言结合为一个有效的"混合系统"等等。或许在某些方面，它不及使用叙词表好，但也有许多叙词表不能比拟的优点，例如成本低，标引速度快，适合用户的检索习惯等。叙词表固然是当今词汇控制的主要方法，但自然语言化已成为一个必然的发展趋势(第17，18章)。

当今世界词表数量在不断增长。为了使采用不同检索语言的情报系统统一、协调地运行，为了使成千上万的数据库能得以"沟通"，词汇的兼容、互换以及多语种的处理已迫在眉睫。现已研究出中介词典、集成词表、宏观词表、微观词表等多种实现词汇兼容与互换的方法。与此相关的算法研究和术语标准化工作也取得了很大的进展(第19，20章)。

(三)

本书是一部难得的情报检索语言专著，有着如下几个特点：

(1) 主次分明，体系得当。这是一部全面论述词汇控制的著作，但作者明确地提出，旨在重点论述叙词表方法。"这是由于在过去20年中，叙词表已成为情报检索中所应用的主要的词汇控制方法"(第1章)。因此，全书有一半以上的篇幅论述的是叙词表方法。就这一点而言，本书是一部关于叙词表编制和使用的简明指南。

(2) 文字精练，图文并茂。本书内容丰富，文字却很精练，该书第一版时有40多万字，第二版的篇幅差不多减少了一半，但却精辟地阐述了词汇控制的各方面问题。全书共有57个图表，还有穿插其中的大量图例，直观性好，便于阅读和理解。

(3) 内容新颖，充分反映新进展。本书全面概述了词汇控制的新方法、新技术。如叙词表的自动编制，混合系统，兼容与互换等都在本书内容之列。充分反映了"新"的特点。

(4) 实用性较强。作为一部专著，本书既具有其理论价值，但也提供了词汇控制的具体程序和方法，诸如如何收集词汇、组织词汇，设计和选择词表结构等都极其具体。

(四)

近年来，我国词表编制和文献数据库建设发展较快，有人称之为"编表热"和"建库热"。"编表热"是随"建库热"而兴起的。据不完全统计，我国已编制并投入使用的叙词表有60多部，而且目前还有一批叙词表在编制之中，其中收词量超过3万以上的大型词表就有五、六部。

我国词表编制工作有如下几个特点：

(1) 发展快。从70年代我国叙词表编制工作开始以来，在10多年的时间里发展迅速，尤其是80年代中后期，图书馆及情报机构竟相编表，投入了大量的人力和物力，编出了一批颇具特色的词表，基本满足了标引与检索工作的需要。

(2) 词表体系初步形成。目前我国正在编制和编成的词表就学科范围而言，有综合性的，有跨学科的，也有专业性的。就规模而言，大、中、小词表都有。各种规模并覆盖各学科的词表体系已初步形成。

(3) 计算机应用于词表编制。例如《社会科学叙词表》、《中国分类主题词表》、《军用主题词表》和《农业科学技术叙词表》等大型词表都运用了计算机编制。计算机在词表编制、管理中得到了越来越广泛的运用。

(4) 词表模式多元化。以《汉语主题词表》为模式的词表结构已不再是我国词表的唯一结构形式，一些新的探索，使我国词表结构呈多元化的趋向。范畴表与词族表合一，字顺表与词族表合一，分类主题一体化等模式已进入我国词表结构之中。

总的来说，我国的词表编制的成就是喜人的。但也毋庸讳言，我国词表编制技术和理论方法研究还不尽理想，和国外相比尚有相当差距。如词表结构、编表技术、机读化、兼容与互换、自然语言的应用等诸多方面还有待加强研究。正是出于这样的考虑，我们将F·W·兰开斯特的《情报检索词汇控制》第二版翻译出来，其目的是希望在我国词表编制工作中能够更多更好地借鉴和吸取80年代国外的新方法和新技术。我们相信，从事编表的人员认真地读一读本书的话，无疑将会从本书中得到一些有益的启示和帮助，从而少走一些弯路。

(五)

我们在翻译过程中力图忠实地反映原书的面貌和风格，保留了绝大多数图表的原样。

尽管F·W·兰开斯特是我国图书情报界熟知的人物，但我们还是根据原书提供的材料，并作了一些补充，撰写了作者简介，目的是让读者们更全面地了解作者，以帮助理解本书。此外，我们还翻译了本书第一版的内容梗概，以展示两版之间的联系和差别。本书的附录是根据Lois Mai Chan 和 Richard Pollard 合编的《联机数据库所用叙词表：分析指南》一书选择其中100部叙词表予以简介，作为了解国外词表的特点和现状的一个窗口，也是帮助读者阅读本书的基础材料。

我们要感谢宝鸡市图书馆陈树年等同志，他们为本书的出版付出了辛勤的劳动。本书有些

章节的翻译还参考了杨劲夫等译的《情报检索词汇规范化》，姚维范、刘昭东译校的《情报系统的兼容性》和北京图书馆图书馆学研究部翻译的《叙词表指南》(英文)，特此向这三本书的译校者致谢。

由于是三人分头翻译，虽经校对和统稿，本书前后翻译上的差异在所难免，加之译者水平有限，不当之处敬请读者批评指正。

译者

1991年3月

本书第一版梗概

1. 检索系统的效率主要取决于存在于该系统中文献类目的大小和组成；而类目的大小和组成则由人们赋予各类目的标记所控制，亦即由描述文献所用的词汇所控制(第 1 章)。

2. 人们赋予某一文献类目的名称(标识)可能是一个词或一个词组，可能是选自某一分类表的类号，也可能是其他符号(例如，任意规定的一个三字符)。类目名称本身对检索系统的性能并无影响。重要的是我们把什么归入某一类目(亦即我们如何确定一个类目的范围)，而不是我们称呼它什么(第 3 章、第 19 章及其他章节)。

3. 受控词汇的作用主要是控制同义词、近义词和同形异义词，将语义相关的一些词联系起来，并提供充分的等级结构以便进行族性检索(第 1 章、第 20 章)。

4. 用于标引的受控词汇必须是合成式的，亦即它必须提供各种组配词汇的手段，以便表达各种文献中论及的任何主题。根据定义，后组式词汇是合成式的(第 2 章)。

5. 分类表、标题表和叙词表都能满足受控词汇的各项要求(第 3 章、第 4 章和第 5 章)。

6. 由于先组系统具有合成的特性，可以表达任何复杂程度的文献主题。但是因为先组系统的文档结构在本质上是线性的，所以该系统不能有效而经济地为文献提供多种存取途径。只有通过提供多重款目和多个参照，先组系统才能提供多种存取途径。但这样做就会形成庞大的检索文档，增加编制和维护索引的费用(第 3 章、第 4 章)。

7. 如果受控词汇具有"文献保证"和"用户保证"，亦即它是根据某个主题领域的文献语言以及根据用户及潜在用户描述他们感兴趣的主题时所用的词汇而编制的，这种词汇可能是最有效的。仅仅通过委员会方式编制的词汇则可能是低效的(第 6 章)。

8. 用于组织和展示词汇的各种方法并不是决定检索系统性能的关键。与如何组织文献类目相比，文献类目的大小和组成则重要得多(第十三章)。尽管如此，对于标引员和检索者来说，某些展示形式可能比另一些展示形式更为有用(第 7 章)。

9. 分面分析对于编制和组织任何类型的受控词汇(包括叙词表在内)，都是非常有用的(第 6、7 章)。

10. 在很多情况下图表展示法颇具吸引力。这种展示法特别适用于采用可视控制台的联

机系统。在美国对图表展示法有所忽视(第 7 章)。

11. 从向标引员和检索者提供最多帮助的观点来看，理想的受控词汇也许是一部叙词表与一部分面分类法的结合。《分面叙词表》即为此类受控词表(第 8 章)。

12. 为了彼此一致并促进各个情报中心和情报机构之间的合作，叙词表的编制和展示必须遵循一定的准则。美国编制叙词表大多采用美国科学技术情报委员会 (COSATI) 的规则(第 9 章)。

13. 叙词表编制的某些规定(例如，关于词的先组程度的规定) 会显著影响一个检索系统的性能，而另一些规定(诸如款目的正装与倒置)则对检索系统性能的影响甚小(第 9 章)。

14. 计算机可以用于：

① 处理和编辑叙词表输入数据，生成反参照，排序和打印出叙词表。

② 有助于叙词表的更新和维护。

③ 存贮机读叙词表以便进行族性检索，自动进行词的转换和替代，并提供有助于标引、检索和词汇控制功能的各种统计(第 11 章)。

15. 受控词汇必须不断更新以反映在文献和提问中出现的新术语(第 12 章)。

16. 词汇的专指性是决定一个检索系统的检准性能的一个主要因素(第 13 章)。

17. 起因于词汇的检索系统的查找失误可能属于下列两种主要类型：

① 缺少专指性而导致的失误；

② 含混和虚假的词间关系而导致的失误(第 13 章)。

18. 像其他任何语言一样，一种完整的情报检索语言通常包括词汇、句法和使用规则(第 14 章)。

19. 标引和检索词汇有三类：叙词(descriptors)、专指词(specifiers)和入口词(entry erms)。这几类词的功能各异，对检索系统性能的影响也各不相同(第 14 章)。

20. 由文献和提问中出现的自然语言语词组成的完整的入口词表，对检索系统的效能和效率是十分重要的(第 14 章)。

21. 为获得高检全率或高检准率(或为同时获得这两者)而设计的一系列手段，提供了检索系统的"句法"(第 15 章)。

22. 在大多数主题领域和大多数文献收藏单位中，不管收藏的规模如何，通常都可以利用

此为试读，需要完整PDF请访问：www.ertongbook.com

最简单的句法有效地在一个系统中进行检索。总的说来，职号对检索系统的性能具有不利的影响。在大多数情况下，词的组配(类目相交)足以减少可能产生的词义模糊(第 16 章)。

23．标引和检索时所用的受控词汇对于手工检索系统来说，实际上是必不可少的。可是不少证据表明：计算机检索系统可以利用文献或文摘中的自然语言有效地运转(第 16 章)。

24．叙词表不用于标引而用作检索的辅助手段，这是可能的(第 16 章)。

25．标引时如采用受控词汇，检索时就不能达到完全的专指。而用自然语言检索则能达到完全的专指(第 16 章)。

26．用自然语言输入，并以叙词表作为检索的辅助手段，两者结合起来可以为标引和检索提供最有用的基础。通过这种结合，用户可以根据个别提问的需求，充分灵活地利用自然语言得到较高的专指度，或利用叙词表进行宽泛的检索(第 16 章)。

27．利用"两级查找法"——不加控制的关键词，加上一个小型的、由宽泛的受控词汇组成的词表，两者结合可使检索非常成功(第 16 章)。

28．利用统计标准以及文献本身包含的词的其他特性，计算机可以进行文献的自动标引(第 17 章)。

29．自动标引可以通过词或词组的"抽取"来实现，也可通过(根据受控词表)"赋词"来实现。后者通常不如前者成功，在任何情况下都不值得采用(第 18 章)。

30．以词同现的统计值为基础，计算机可被用于词的分类，以便自动形成词的类目。这些类目在检索实践中具有潜在的价值；在某些情况下它们能替代人工编制的常规叙词表（第 17 章）。

31．为了获得最高效能，必须根据特定的机构或用户群体的需求精心编制受控词表（第 18 章）。

32．日益频繁的机构间的合作以及日益增加的文献加工量，迫使人们研制各种实现受控词汇之间兼容和互换的方法。下列方法可用于词汇的互换：利用共同的范畴表或中介词典；利用计算机将一种受控词汇自动转换成另一种受控词汇；精心编制微观叙词表，使之适合于某些大型叙词表中的等级结构(第 18 章)。

33．一种情报检索语言应该具有下列功能：

① 使标引员前后一致地表达文献的主题概念；

② 使检索者使用的词汇与标引员使用的词汇一致；

③ 提供各种手段，使检索者得以视实际情况的不同需求，改变检索策略，借以得到较高的检全率或检准率(第 20 章)。

34. 如果一部受控词表中词汇的专指度不高，并出现不明确或不正确的词间关系，可以造成系统失误(第 13 章)。

35. 一部受控词表可能是指定性的或建议性的，也可能是两者兼而有之。词表的指定性越强，标引结果就越可能一致(第 20 章)。

36. 提供各类辅助工具是可能的。它们包括：预印的词汇一览表、各类入口词表、词汇变动历史的文档、词频统计以及检索策略的文档(第 21 章)。

37. 在一个大型的多功能的情报系统中，需要研制出将一套词汇的自动转换为另一套词汇（例如，为了出版的目的）的方法。为了不断更新词汇，还可能需要一些相当复杂的方法(第 22 章)。

38. 联机检索系统可以为词汇的显示、词表的维护以及标引、检索时对词汇的选用提供新的方法(第 23 章)。

39. 一部非常完整的入口词表，对使用非代表性查找模式的联机检索系统，可能特别重要(第 23 章)。

40. 联机模式显得特别适合于自然语言检索系统(第 23 章)。

41. 编制和维护受控词表的费用较为昂贵。根据成本-效益分析，采用一些比较简单的控制词汇的方法，可能是正确的(第 24 章)。

目　录

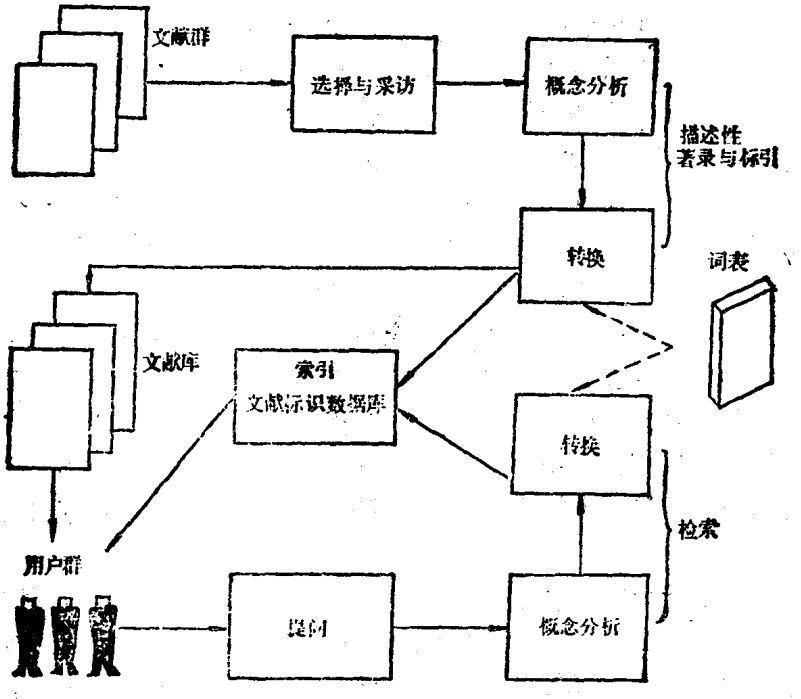
图表目录

第 1 章　为什么要控制词汇

在情报检索系统中，通常对用于描述所论文献主题的词汇需要进行控制。情报检索词汇控制（如本书的书名所示）讨论情报检索中词汇控制各方面的问题。本书特别着重论述叙词表，这是由于在过去 20 年中，叙词表已成为情报检索中所应用的主要的词汇控制方法。尽管如此，本书也论及其他方法，其中包括情报检索系统运行时词汇不加控制。

典型的情报检索系统的主要组成部分如图表 1 所示。输入部分由操作检索系统的情报中心所采集得来的大量文献（此词含义极广，包括印刷型以及其他所有类型的知识记录）组成。这表示存在着选择文献的各种准则和方针，因而也表示要详细而准确地了解所服务的组织机构的情报需求。文献一经采得，就需要加以“组织和控制”，使它便于识别和查找，以满足用户的各类需求。组织和控制手段包括分类、编目、主题标引和撰写文摘。对文献外部形态的描述（描述性编目）和检索点（例如：著者、题名）的选取，是组织和控制文献的两个重要措施，它们可使目录和书目中的记录便于检索。

如图表 1 所示，主题标引过程包括两个显然不同的智力步骤：文献的“概念分析”和将概念分析的结果“转换”（translation）成特定词表中的词汇。要有效地进行概念分析，标引员必须理解文献的主题，同时应充分了解系统用户的需求。



图表 1　情报检索系统的主要组成部分

标引过程的第二步骤是将概念分析的结果转换成特定词表中的词汇。在大多数系统中，此步都涉及一个“受控词表”，即用于表达文献主题的有限的词汇集合。这种词表可以是一部

此为试读，需要完整PDF请访问：www.ertongbook.com