

美国DIALOG

国际联机检索方法

张惠惠 编

上海交通大学出版社

DIALOG



美国 DIALOG 国际联机 检索方法

张惠惠 编

上海交通大学出版社

内 容 简 介

本书主要介绍美国 DIALOG 联机检索系统使用方法。全书共分 DIALOG 系统介绍、检索项、检索逻辑、DIALOG 系统指令、检索输出、检索实例、DIALOG 系统其他检索功能和服务、DIALOG 系统数据库专业介绍、附录等八章。

本书可作为广大图书馆或情报工作人员、科技干部和其他工程技术人员自学使用美国 DIALOG 联机检索的参考书,也可作为从事美国 DIALOG 联机检索工作人员的必备使用手册或培训教材。

美国 DIALOG

国际联机检索方法

出 版: 上海交通大学出版社
(淮海中路1984弄19号)

排 版: 浙江上虞汤浦印刷厂
发 行: 新华书店上海发行所
印 刷: 江苏太仓印刷厂
开 本: 787×1092(毫米) 1/16

印 张: 15.75

字 数: 384000

版 次: 1989年4月 第1版

印 次: 1989年4月 第1次

印 数: 1—3300

科 目: 187—305

ISBN7-313-00361-7/TP·39

定 价: 5.10元

前 言

随着电脑和远距离通信技术的出现及两者的结合,使情报的收集、贮存、处理和传递技术进入了一个新的阶段,国际联机情报检索技术也就此应运而生。

当前,国际联机情报检索技术的应用相当普遍,目前较大的联机检索系统有 100 多个,著名的有 DIALOG、ORBIT、ESA/IRS 和 MEDLINE、BRS、JOIS 等系统,其中以美国 DIALOG 联机检索系统为最大。

全书根据 DIALOG 最新资料编著而成,内容有 DIALOG 系统介绍、检索项、检索逻辑、DIALOG 系统指令、检索输出、检索实例、DIALOG 系统其他检索功能和服务、DIALOG 系统数据库专业介绍、附录等八章,对整个 DIALOG 系统的功能及操作方法作了较全面、详细的阐述,对使用 DIALOG 系统的检索人员、科技干部和其他工程技术人员是一本较好的使用指南。

本书在编写过程中曾得到上海交通大学图书馆和上海交通大学情报科学技术研究所吴善勤教授、何大镛和杨宗英副教授、葛玉龙副研究员的大力支持、帮助和审读,上海交通大学科技外语系吴银庚教授的翻译校对,在此,致以衷心的感谢。

限于作者水平,书中错误难免,望请读者批评、指正。

编 者

1988 年 12 月

目 录

第一章 DIALOG 系统介绍	1
1-1 DIALOG 系统检索	2
1-2 检索准备	2
1-3 联机检索过程	2
1-4 DIALOG 系统数据库介绍	2
1. 文档	2
2. 记录	3
3. 字段	3
4. 记录标引	3
5. 数据库索引	3
6. 蓝页	5
7. 联机词表	5
第二章 检索项	7
2-1 集号	7
2-2 单元词	7
2-3 词组	7
2-4 截断	8
1. 开放式截断	8
2. 限制式截断	8
3. 嵌入式截断	9
4. 说明	9
2-5 标点符	9
1. 单元词中的标点符	9
2. 词组中的标点符	10
3. EXPAND 指令中的标点符	10
2-6 检索项目的限制	11
1. 基本索引和辅助索引	11
2. 其他限制代码	11
3. 后缀代码	11
4. 前缀代码	13
5. 限制后缀代码	13
6. 限制时间周期	14
第三章 检索逻辑	16
3-1 OR 逻辑组配	16

1. OR 运算	16
2. 截断检索	16
3. 范围检索	16
3-2 AND 逻辑组配	17
1. AND 运算	17
2. DIALOG 系统的邻近算符	17
3-3 NOT 逻辑组配	19
1. NOT 运算	19
2. NOT 邻近算符	19
3-4 SELECT 语句中使用多个算符	20
1. 处理顺序	20
2. 说明	20
第四章 DIALOG 系统指令	22
4-1 DIALOG 指令一般格式	22
4-2 DIALOG 指令一览表	22
4-3 DIALOG 指令及其用法	23
1. BEGIN 指令	23
2. SELECT 指令	25
3. SELECT FILES 指令	28
4. EXPAND 指令	30
5. PAGE 指令	36
6. PAGE-指令	37
7. DISPLAY 指令	37
8. TYPE 指令	40
9. PRINT 指令	42
10. PRINT VIA DIALMAIL 指令	45
11. PRINT TITLE 指令	46
12. PRINT QUERY 指令	47
13. PRINT QUERY ACTIVE 指令	48
14. PRINT QUERY DETAIL 指令	48
15. PRINT QUERY ACTIVE DETAIL 指令	48
16. PRINT CANCEL 指令	48
17. LOGOFF 指令	49
18. LOGOFF HOLD 指令	49
19. FILE 指令	50
20. EXPLAN 指令	52
21. LIMITALL 指令	53
22. SAVE 指令	54
23. MAP 指令	59

24. RECALL 指令	65
25. RELEASE 指令	67
26. REPORT 指令	68
27. EXECUTE 指令	73
28. KEEP 指令	76
29. KEEP CANCEL 指令	77
30. ORDER 指令	77
31. ORDER ITEM 指令	78
32. ORDER ADDRESS 指令	79
33. ORDER CANCEL 指令	80
34. ORDER LIST 指令	80
35. ORDER REVIEW 指令	82
36. SET 指令	83
37. SHOW 指令	87
38. SORT 指令	90
39. EDIT 指令	93
40. COST 指令	102
第五章 检索输出	103
5-1 编号格式	103
5-2 用户规定格式	103
5-3 其他输出格式	104
1. TAG 输出格式	104
2. KWIC 输出格式	106
3. HILIGHT 输出格式	108
4. SORT 输出格式	109
5. REPORT 输出格式	109
6. PRINT 输出编排格式	111
7. PRINT VIA DIALMAIL 输出格式	111
8. NOHEADER 输出格式	111
第六章 检索实例	112
第七章 DIALOG 系统其他检索功能和服务	117
7-1 一次性检索功能	117
1. 一次性检索的文档联接	117
2. 一次性检索的检索项	117
3. SET 运算	118
4. FROM 运算	120
7-2 贮存检索功能	122
1. SAVE SDI 功能	122
2. MAP 功能	124

7-3	DIALOG 检索通讯软件(DIALOGLINK™)	128
1.	DIALOGLINK 软件功能	128
2.	DIALOGLINK 使用方法	128
7-4	DIALOG 光盘检索(DIALOG OnDisc™)	129
1.	DIALOG 系统 ERIC 光盘 (DIALOG OnDisc™ ERIC)	130
2.	DIALOG 系统 MEDLINE 光盘 (DIALOG OnDisc™ MEDLINE®)	131
3.	DIALOG 系统 NTIS 光盘 (DIALOG OnDisc™ NTIS)	131
4.	DIALOG 系统 STANDARD & POOR'S CORPORATIONS 光盘 (DIALOG OnDisc™ STANDARD & POOR'S CORPORATIONS)	132
7-5	DIALOG Business Connection SM (DBC)	132
7-6	DIALOG Medical Connection SM (DMC)	133
7-7	DIALINDEX	134
1.	功能	134
2.	使用方法	135
7-8	DIALORDER 联机订购	136
7-9	DIALMAIL 电子邮件服务	137
7-10	DIALOG Gateways (接口通道)	139
1.	OAG ELECTRONIC EDITION	139
2.	DIALOG QUOTES AND TRADING	139
第八章	DIALOG 系统数据库专业介绍	141
附录一	DIALOG 系统功能范围	219
附录二	DIALOG 系统标点、符号和检索词	221
附录三	MAP 贮存代码和使用文档	223
附录四	DIALOG 数据库 一次性检索分类目录表	225
附录五	联机检索工作单	241

第一章 DIALOG系统介绍

美国 DIALOG 情报服务公司是当今世界上最大的联机情报检索系统,总部设在美国加利福尼亚州旧金山附近的帕洛·阿尔托市(Palo Alto)。该系统从 1972 年起服务至 1988 年 6 月已有 300 多个数据库,专业内容有:社会科学、化学、物理、数学、医学、材料学、时事、建筑学、教育、经济学、商业、公司、人文学、宇航、能源、法律、工程、农业、计算机、生物等。数据库文献类型有:杂志、会议录、图书、论文、进展中科技报告、政府文件、专利、标准、市场行情、经济预测、统计数据、名人传记等。

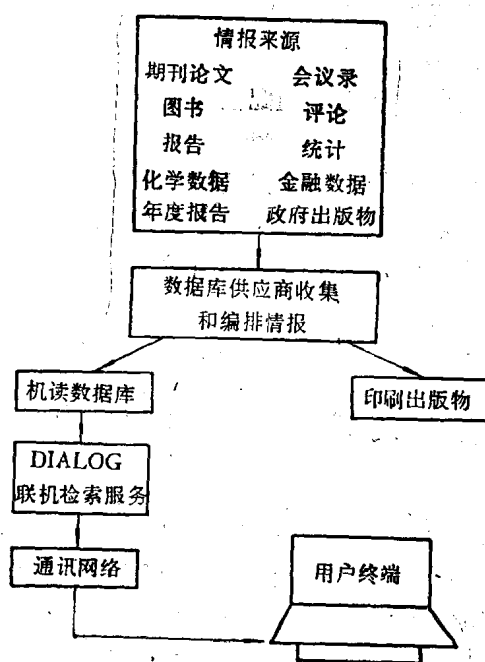
DIALOG 数据库由数据库供应商(主要由私人、国家或政府机构提供的出版物)提供,并以机读形式发给 DIALOG。DIALOG 系统对每篇记录内容进行检索标引,提供索引和原文入藏。

目前,DIALOG 数据库存有 15.2 亿篇以上的记录,记录类型有:文摘的文献型记录;公司、机构或个人的目录型记录;统计或金融情报的表格型记录;章节或文章的全文型记录。

DIALOG 系统主机设在加利福尼亚帕洛·阿尔托市的 DIALOG 情报服务公司。计算机由运行检索 DIALOG 软件的中央处理机、电讯部分、带有数据库和索引的磁盘驱动器三部分组成。

中央处理机由四个独立系统(A、B、C、D)组成,用户用哪一个中央处理机取决于所用的通讯网络。

联机流程图



1-1 DIALOG 系统检索

在 DIALOG 系统上进行检索,主要是通过用户与计算机之间的对话。当终端上出现系统提示(出现问号?)时,用户输入指令,告诉计算机要做什么,并按回车键(ENTER 键)将指令发给 DIALOG,系统对用户发出指令进行响应(例如,查找每个检索项并显示命中数),再出现新的提示。在一次检索中,用户可检索同一内容不同数据库,同一数据库的不同内容,不同内容的不同数据库,也可一次检索多个文档,也可灵活使用各种指令(详见有关章节)。

1-2 检索准备

由于 DIALOG 是一个指令驱动检索过程,所以在进行检索前要做好下述几项准备。

(1) 详细确定要查找的情报,并尽可能进行完整性的描述。

(2) 标引所要查找情报的不同概念(即课题的每一个部分),千万不要遗漏,否则会改变课题内容。

(3) 对每一个概念,确定是否有同义,或单、复数,或英、美不同拼音,或者是具有相同含义的完全不同的词,把这些概念组成所要查找的课题。

上述三项都要认真填入工作单(工作单见附录)。

(4) 确定数据库(可从 DIALOG Database Catalog, Database Bluesheets, DIALINDEX 中去确定)。

1-3 联机检索过程

与 DIALOG 联机后,基本联机检索有四步:

(1) 用 BEGIN n (n 为文档号)指令联接规定数据库。例:BEGIN 41 或 B 41。

(2) 用 SELECT(或 S)指令进行查找、组配检索项。例:S CANADA ; S RIAN OR SNOW,S S2 ANDACID

(3) 用 TYPE 指令、DISPLAY 指令联机显示检索结果。例:T S1/5/1。

用 PRINT 指令脱机打印检索结果。例:PR S1/5/ALL。

(4) 用 BEGIN 指令改变数据库,或用 LOGOFF 指令与 DIALOG 设备断开。

(上述指令用法详见有关章节)

1-4 DIALOG 系统数据库介绍

1. 文 档

一个单独的数据库也称为文档,并给定一个文档号。大的数据库按年份分成若干个文档,分别给予单独的文档号;每个文档可单独检索。DIALOG Database Catalog 对每个数据库有一简要的介绍。DIALINDEX 是各文档的分类索引,这些都供用户在联机前选择数据库之用。

2. 记 录

DIALOG 数据库的信息项目称为记录。它有以下几种类型:

- (1) 文献型(带或不带文摘)——科技资料文档;
- (2) 目录型——公司机构或个人文档;
- (3) 统计表格型——统计及金融文档;
- (4) 全文型——新闻及部分经济文档。

每篇记录都编有一个按顺序排列的 DIALOG 入藏号。记录中每一项目均可检索。

3. 字 段

每个记录分成若干片断或字段,用于检索或显示。在文献型记录中有如下字段:

- ①文献题目, ②文献作者, ③文献出版日期, ④机构, ⑤语种, ⑥文献出处, ⑦文摘, ⑧叙词, ⑨标引词, ⑩分类, ⑪节标题等。

在非文献型记录中,还有公司或机构名,公司的净销售或总销售,公司所在城市、电话、邮政编码、产品名、全文等。

4. 记 录 标 引

记录中每个字段内容都有一种检索标引。标引有词标引、词组标引、数字标引。有些字段用词或词组标引两次。数据库蓝页标有详细字段说明。

- (1) 词标引 词标引是一系列字母或数字符号,周围有标点或空格,例:X-RAY,1,100。通常有九个禁词不作词标引:AN、AND、BY、FOR、FROM、OF、THE、TO、WITH。

- (2) 词组标引 通常包括标点和空格。

例:作者标引:Dickenson, Anne

Olds, Henry F., Jr.

期刊名标引:Telephone Engineer & Mgmt

- (3) 词和词组标引 有些词在词和词组中标引两次。DIALOG 系统中描述主题的有叙词字段(由数据库供应商提供的词表)和标引词字段(非词表词)。但标引词在文献中出现,也由数据库供应商提供。一个标引词可以是很专、很新,以至最后编入词表。

例:叙词标引:computer data processing
 computer design design
 data processing

5. 数 据 库 索 引

每个数据库都有两种索引——基本索引和辅助索引。

- (1) 基本索引(后缀字段限制) 由文献篇名、文摘和主题词等字段构成(不包括 AN, AND, BY, FOR, FROM, OF, THE, TO, WITH 九个虚词)。

索引中的词一般是以单元词形式在文档中按字顺排列。每个词后面都带有计算机可识别的文献记录存取号和查找地址。每个数据库要求后缀代码不同。

基本索引后缀代码:

...../AB 文摘字段

...../DE,...../DE* 叙词字段(单词或多元词主题项)

...../DF,...../DF* 所有叙词(单元词)

...../ID,...../ID* 标识字段(一种半控制自由语言项,比规定叙词词汇表要多,包括单元词或多元词,为得到更好检索,一般采用全文检索方法)。

...../IF,...../IF* 所有标识(单元词)

...../TI 标题(检索标题中有定义的单词,包括副标题、主标题和相关叙词内容)。

...../NT 说明

...../SH 标题

说明

* 表示在一些数据库中为主要的。

AB—ABSTRACT

DE—DESCRIPTION

ID—IDENTIFIER

TI—TITLE

NT—NOTE

SH—SECTION HEADING

(2) 辅助索引 该索引通常为非主题性索引。它由独立的词或完整的字符串构成。各个文档中辅助索引的数量不尽相同。

辅助索引的基本种类有:

AU = 著者(Author)

AN = 文摘号(Abstract Number)

CC = 分类代码(Class Code)

CS = 机构来源(Corporate Source)

DT = 文献类型(Document Type)

PA = 专利代理人(Patent Assignee)

PN = 专利号(Patent Number)

JN = 期刊名(Journal Name)

LA = 文种(Language)

RN = 报告号(Report Number)

PY = 出版年份(Publication year)

UD = 更新日期(Update)

JC = 杂志代码(Journal Code)

DC = 标识代码(Descriptor Code)

PC = 产品分类(Product Classification)

EC = 销售事态分类代码(Event Codes)

必须注意: 由于各文档的专业及文献类型不同,前缀代码表示也有所不同。例:在“化学文档”(CHEMSIS)中 RN = 表示化学物质登记号,而在科学文摘中(INSPEC)RN = 表示报告号。

所以,在联机检索时,应首先看一下文档的蓝页介绍。

6. 蓝 页

这是 DIALOG 每个数据库的内容和文档结构的描述,包括文档号、文档说明、主题范围、数据库资料来源、文档规模、收录时间、更新周期、累计篇数、检索实例、基本索引、辅助索引、限制代码、联机分类代码、MAP 贮存代码、REPORT 代码、输出格式、DIALOG 入藏号输出、联机输出功能等。

7. 联 机 词 表

指令功能:	EXPAND	} (指令详细说明见有关章
	PAGE	
	PAGE-	
	SELECT	

DIALOG 一些数据库提供的联机词表(一种主题词字典)属于联机文档的一部分。词表中主题词通常作为叙词(descriptors-DE)标引。词表中显示完整的词组,包括标点符号和空格(其余为单元词)。

词表中的主题词往往是代表数据库文献的词,用户使用该词可以了解用该词标引的文献有多少。

分层的联机词表标引主题的级项,对某一词标引其上位类、下位类、参考项及代用项,对非分层的词表,可标引专业标引项,也可以为同义词项,用? THESAURI 指令,可联机查询数据库有否词表及词表信息。

例如下面为一-ERIC 叙词词表。

PRIVATE SCHOOLS	Jul.1966
CIJE:748 RIE:856	GC:340
UF Independent Schools	
Nonpublic Schools	
NT Parochial Schools	
Private Colleges	
Proprietary Schools	
BT Schools	
RT Correspondence Schools	
Private Education	
Private School Aid	
Single Sex Schools	

用 EXPAND 指令,后跟具体项目扩展该项词表。

例 1.

? expand private schools

Ref Items RT Index-term

E1 408 19 PRIVATE SCHOOL AID

E2 1 PRIVATE SCHOOL RELATIONSHIP

E3 1952 10 *PRIVATE SCHOOLS

E4 141 PRIVATE SECTOR

⋮ ⋮ ⋮ ⋮ ⋮

E12 1 *PRIVATED

Enter P or E for more

用 EXPAND 指令后跟 E 号,扩展该 E 号参考项,列出所有与该 E 号相关的项,在该扩展表中 TYPE 一列中列出相关项表示与 E 项之间的关系。

例 2.

? expand e3

Ref	Items	Type	RT	Index-term
R1	1952		10	*PRIVATE SCHOOLS
R2	0	U	1	INDEPENDENT SCHOOLS
R3	0	U	1	NONPUBLIC SCHOOLS
R4	596	N	8	PAROCHIAL SCHOOLS
R5	2318	N	12	PRIVATE COLLEGES
R6	293	N	9	PROPRIETARY SCHOOLS
R7	83786	B	114	SCHOOLS
R8	112	R	8	CORRESPONDENCE SCHOOLS
R9	261	R	14	PRIVATE EDUCATION
R10	408	R	19	PRIVATE SCHOOL AID
R11	28	R	8	SINGLE SEX SCHOOLS

? select r1 Or r5

1952 PRIVATE SCHOOLS

2318 PRIVATE COLLEGES

S1 4224 "PRIVATE SCHOOLS" OR "PRIVATE COLLEGES"

F——取被扩展项,不用扩展项;

U——用扩展项;

N——下位项;

B——上位项;

R——参考项(既不属上位,又不属下位),对参考项可以再扩展。

第二章 检索项

检索项及其组成的表达式用来确定联机检索课题。DIALOG 系统在用户连接的数据库中,用 SELECT 指令在基本索引和辅助索引中查找检索项。用户也可先用 EXPAND 指令扩展数据库字典,然后用 SELECT 指令选择正确的检索项(详见 SELECT 指令及 EXPAND 指令章节)。

检索项可以是单元词、词组,其中可以是字母、数字或字母数字的组合。单元词表达较简单,但词组表达常有多种方法,同时在检索时必须小心。

2-1 集 号

DIALOG 系统对检索项进行 SELECT 指令操作后,给每个检索项赋予一个集号,用字母 S 后跟编号表示(例: S1)。SELECT 指令只产生最后集号,SELECT STEPS 指令对每一项产生一个集号(详见 SELECT 指令章节)。

2-2 单 元 词

任何一个单词、数字或缩写表达式都可作为一个单元词输入到 SELECT 语句中(除去九个禁词:AN、AND、BY、FOR、FROM、OF、THE、TO、WITH)。SELECT 语句中的标点符号(例:连字符、撇号)必须用(W)运算符表示。

例: FREQUENCIES VIDEO

IBM B 12

100 V 8

(1) 有些词或字母数字组合作为单元词放在 SELECT 语句中必须加引号。

例: 'AND'、'NOT'、'FROM'、'OR'、'S 3'、'E 12'、'R 6'等。

(2) 多个单元词之间可以用逻辑符和邻近符连接。

2-3 词 组

DIALOG 系统的数据库基本索引中由单元词和词组两种表达形式。在 SELECT 语句中,词组是作为一个完整短语(包括标点)来进行查找。当用缩写字母时,必须精确输入,以免误检。对非规范化的词组,各单词间要用邻近算符连接。

例: 非规范化词组: RIO DE JANEIRO

检索表达式: S RIO(W)DE(W)JANEIRO

(1) 在数据库的基本索引中,词组一般在叙词(规范化词)和标引词(非规范化)字段中查找(见基本索引有关内容)。

例: CAD/CAM

COMPUTER-AIDED DESIGN

VITAMIN B₁₂

X RAYS

(2) 同样,在词组中遇到 AND、NOT、FROM、OR、S₃、E₁₂、R₆ 等词也必须加引号。

(3) 在辅助索引中词组作为整体输入。

例: CO=HUTTON(E,F.)GROUP INC. ——公司名字段

2-4 截 断

DIALOG 系统检索是根据检索项与索引的实际匹配情况进行的。有时用户要检索一个词的不同形式,例:单、复数名词或动词的过去时和现在时。截断就是检索一个词的不同形式,或英美不同拼音,这是一种 OR 的逻辑选择允许的不同形式。

截断就是利用某些检索词的词干或不完整词形进行检索,使所用检索词包括了该词词尾的多种变化形式或词后跟着其他字的任意词,从而达到较高的查全率。

1. 开放式截断

(1) 词干后跟任何个字母,用 ? 表示。

例: ? S DYE?

检索: DYE

DYEABILITY

DYEBATH

DYEFONATE

DYES 等。

(2) 在词组的任何地方放一个问号,那么问号处可出现字母、数字、标点符或空格。

(3) 检索作者时要注意:

SS AU=GREEN,?

表示检索姓为 GREEN 的其他名的人。

SS AU=GREEN?

表示检索以 GREEN 为词干的所有姓的人。

2. 限制式截断

(1) 词干后只跟一个字母,用 ?? 表示,两问号间空一格。该截断主要用于单、复数名词的检索。

例: S DYE??

表示检索 DYE,DYED,DYES。

(2) 词干后至多跟与问号数相同的字母数,用 ?? 或 ??? 表示,问号间不用空格。

例: SS UNIVERS??

表示最多跟两个字母,可检索 UNIVERSE、UNIVERSAL 等。

SS CAND???

表示最多跟三个字母,可检索 CANDY、CANDIES、CANDIED 等。

3. 嵌入式截断

该截断一般用在单词中间,嵌入字母数等于问号数,通常用在英美不同拼法。

例: SS WOM ? N ?

表示检索 WOMAN, WOMANLY, WOMEN。

SS FIB ?? BOARD

表示检索 FIBERBOARD, FIBREBOARD。

4. 说 明

(1) 使用截断必须慎重,词干不要太短,以免检出许多与原来含义不相关的项目。

例: 以 CAT 为词干的词许许多多,其含义各不相同,有 CATALOG, CATASTROPHE, CATHEDRAL 等。

又如以 COM 为词干的词也有同样情况。所以,对这类词使用截断必须小心,可输入较长词干或全部词输入,以免误检。

(2) 英美不同拼音的词,如变化字母数不同则不能用嵌入式截断检索,必须详细写出并用 OR 组配后输入。

例: S COLOR OR COLOUR

(3) EXPAND 指令中的检索项不能用截断。

(4) DIALOG 系统一次搜索 20000 个条目,检索以该词干标引的 2000 个单词,当超过 20000 个条目和 2000 个单词时,系统发出一个信息,标引最后一个处理项目,并标引前 2000 个条目或单词,形成一个集。这个集不表示该截断项目的完整检索。通常要修改检索策略,并用其它方法检索截断项目。

2-5 标 点 符

1. 单元词中的标点符

SELECT 语句中使用带有标点符的单元词,该标点符必须用(W)运算符代替。由于标点符是检索项的不完整分隔,不用(W)时,则标点符两边作两个词处理。例: ALZHEIMER'S 作为 ALZHEIMER 与 S 两个词,所以必须用 SELECT ALZHEIMER(W)S。

例: ? select alzheimers

S 1 0 ALZHEIMERS

? select alzheimer(w)s

37 ALZHEIMER

2132 S

S 2 35 ALZHEIMER(W)S

? type S 2/6/1

2/6/1

0017165