

随着现代科学技术的飞速发展, 科学研究领域不断扩大, 科技文献的品种和数量爆炸性地增长, 微生物学当不例外。尤其是微生物学涉及的面很广, 它与分子生物学、生物化学、免疫化学、医学、病毒学和遗传学等学科都有密切的关系; 在工业、农业、轻工业、医药、环保和国防中有广泛应用。因此微生物学文献存在严重的分散性, 分散在许多邻近学科期刊中。这给传统的手工搜集情报资料、查阅文献带来很大困难。利用晚近发展起来的国际联机情报检索 (International On-Line Information Retrieval) 可以较好的克服这个问题。

(一) 国际联机检索的特点

国际联机检索是 70—80 年代出现的一种新的现代化情报检索手段。它是利用检索终端, 通过卫星通讯网络与地球上任何地方与之相联系的情报检索系统联机, 运用一些特定的指令和检索策略与检索系统的电脑进行“人机对话”, 从预先编制并存储了的数据库中查找所需情报资料和文献的过程。具有检索范围广泛而深入、快速、高效的特点。

1. 检索速度快, 效率高: 因为电脑具有高速处理信息的功能 (每秒钟可处理上千万条指令), 所以联机检索仅用几分钟就可以查遍几年、几十年的有关情报资料, 可提供世界各国几万种期刊中的有关文献, 这是以手工方式检索无法比拟的。

2. 检索范围广: 由于联机是用卫星通讯网络把检索终端和检索系统联系起来的, 所以它不受国家和地理位置的限制, 可以跨越地区或国家进行检索。任何读者可以通过检索终端与远隔重洋的、存储有数以百万计的各种文献资料的数据库中查找自己所需要的文献, 而且立刻可以在检索终端显示出来。

3. 检索深度深: 联机检索不但范围广泛, 而且深度也深。表现在内容新, 查准和查全率高。检索系统的数据库一般更新很快, 更新周期最慢一个季度、一个月, 快的一周。因此利用联机能检索到国际最新文献。一个熟练的手工检索者, 查到文献容易, 但要查全是相当困难的, 漏检率很高。联机检索则可避免漏检。因为检索系统的容量很大, 数据库又多, 每一个数据库的存储量也很大, 其存储的文献少则几万篇, 多的几百万篇。由于电脑具有高速的功能, 可以按照确定的检索策略到多个数据库中去查找, 因而可以把交叉文档中的有关文献全部检索出来, 这是手检难以做到的。还可以根据检索需要, 确定以主题、作者、出版时间或分

类代码等为检索途径进行检索, 并随时可以修改检索策略, 保证了联机检索较高的准确率。

(二) 国际联机的结构

国际联机检索系统的结构如图 1 所示:



图 1 国际联机检索系统的结构示意图

1. 检索终端: 是检索者与检索系统的主机电脑进行联系的装置, 有电传终端、打印终端和带显示屏的视频终端等。检索者可以通过终端上的键盘, 用电脑所能识别的特定指令和检索策略与主机电脑进行“人机对话”, 不但可以提问, 而且可以修改提问, 直到获得满意的检索结果。检索结果可由终端直接打印输出 (费用较高), 也可要求系统脱机打印 (费用较低), 迅速航寄给终端用户, 还可通过终端向检索系统订购原文的复制件。终端与检索系统的主机电脑之间的联系是通过通讯网络来实现的。

2. 主机电脑: 这是联机检索系统的核心部分, 它负责整个系统的运行管理。各个检索系统选用不同类型的电脑, 要求其容量大、速度快和功能强。

3. 数据库: 是联机检索系统的存储部分, 也是检索系统的基础。它是由若干个有关联的文档 (File) 所组成的集合体, 存储着大量有价值的文献记录和数据。内容分为两类: 文献型数据库 (Database) 和数值型数据库 (Databank)。前者由文献记录组成, 后者由具体的数据和图表等组成。

4. 卫星通讯网络: 这是联机检索系统的联络手段, 以电波传递着信息。

(三) 国际联机检索系统

目前国际联机情报检索已发展到相当高的水平, 其网络和终端已遍布世界主要国家和地区。1980 年中国建筑科学研究院等单位联合在香港装设了我国第一台国际联机检索终端。现在联机检索终端已普及到北京、上海、许多省会城市和大型图书情报单位, 可以提供联机情报检索服务。据统计全世界已有数据库 1100 多种, 较大的国际联机检索系统也有 100 多个。最大的是美国洛克希德 (Lockhead) 公司的 DIALOG 系统^[1,2], 其次是美国系统开发公司 (SDC) 的 ORBIT 系

统^[3],美国国家医学图书馆的 MEDLINE 系统,还有欧洲空间组织的 ESA/IRS 系统^[4]和日本国家科技情报中心的 JOIS 系统等。下面重点介绍 DIALOG 系统。

DIALOG 系统设在美国加利福尼亚州的帕洛·阿尔托 (Palo. Alto) 市。它拥有两台大型的电子计算机,其中一台与 TYMNET 卫星通讯网络联线,另一台与 TELENET 卫星通讯网络联线。总运算能力为 14 MIPS,即每秒钟可处理 1400 万条指令。现该系统的数据库有 300 多个文档,共存储文献超过 1 亿篇,占世界机存文献总量的 50% 以上。其中包括 40 多个语种,5 万多种期刊,占世界现有发行期刊总量的 60%。文档的专业包括:综合性学科、自然科学、应用科学和工艺学等。收集的文献类型有:书报、刊物、学位论文、专利、会议录和政府工作报告等。特别应该指出该系统具有一个 SSI 数据库,存储当前世界上正在进行研究的课题及近两年内完成的课题。这是其他出版物所不见到的信息,为科研工作者提供了当前国际攻克新课题和研制新产品的重要线索。还有该系统新增加了一个文档 WPI 和 WPIL 世界专利索引,这对查找世界专利文献提供了一条简便和快速的有效检索途径。DIALOG 系统每周工作五天,每天 22 小时。现拥有终端用户 5 万多个,遍及全世界几十个国家 and 地区的 200 多个城市。检索有关微生物学文献的数据库及文档号可参照附录。

(四) 国际联机的检索方式

国际联机情报检索系统,主要有以下两种检索方式

1. 定题检索 (Selective Dissemination of Information, 简称 SDI): 相当于查找现刊情报资料。它是根据读者的情报需求,预定好检索内容,长期保留在检索系统的主机电脑中,定期一个月、半个月或一周从新到数

.....

Welcome to DIALOG

?b 55---(1)

O4 nov 87 14:36:35 User 068217---(2)

\$0.15 0.005 Hrs File1---(3)

File 55:BIOSIS PREVIEWMS 81-87/OCT BA8408; RRM3308---(4)

Set Items Description---(5)

?ss raw (f) starch/ti, de---(6)

ss raw (f) starch/ti, de---(7)

S1 1637 RAW/TI, DE---(8)

S2 5341 STARCH/TI, DE

S3 59 RAW (f) STARCH/TI, DE

?ss (digesti??* enzyme)/ti, de + (ethanol + alcohol) (f) fermentation/ti, de + gluco() amylase/ti, de---(9)

ss (digesti??* enzyme) ti, de + (ethanol + alcohol) (f) fermentation/ti, de + gluco() amylase/ti, de---(10)

S4 8208 DIGESTI??/TI, DE

S5 53941 ENZYME/TI, DE

数据库的情报资料中检出有关文献,提供给读者。这种检索对掌握某一学科或同类专题的最新水平和发展动态特别合适。

2. 追溯检索 (Retrospective Search, 简称 RS): 相当于普查以往的情报资料。不仅可以查找最新资料,而且可以根据读者的要求从需求情报时起,追溯到过去某时间止,作一次回溯查找全部或部分文献资料。因此一次检索就能掌握相当长时期以来所积累起来的全部文献资料。这种检索特别适合于申请专利、撰写专题评论以及从事新课题研究,需要全面系统掌握有关文献资料的要求。

(五) 国际联机检索微生物学文献举例

国际联机检索微生物学文献,是用用户终端与 DIALOG 系统进行人机对话,通过系统的指令查找有关数据库记录中所标引的关键词或主题词来实现的。首先要明确需要检索的课题,并进行细致的分析。其次选定适当的数据库文档号。第三是确定合理的检索策略,输入关键词,并进行恰当的逻辑匹配和必要的反馈与修正。读者可以自己上机检索,也可以填写《联机检索提问单》委托终端操作人员代为检索。

下面是利用 DIALOG 系统的一个检索实例(部分)及解析,以资参考。

课题: 生淀粉酒精发酵。

关键词: 生淀粉 (Raw Starch); 酒精发酵 (Ethanol Fermentation; Alcohol Ferm.)。注意关键词的同义词,如 Ethanol 与 Alcohol 同义,为了避免漏检,应把有关同义词输入主机查找。

调用数据库文档 55 号 (Biosis Previews)。

当终端通过卫星通讯网络与主机联机、显象屏上出现 "Welcome to DIALOG" (欢迎使用 DIALOG) 字样时,即可输入指令开始检索。

S6 11435 ETHANOL/TI, DE
 S7 13067 ALCOHOL/TI, DE
 S8 6106 FERMENTATION/TI, DE
 S9 531 (ETHANOL/TI, DE + ALCOHOL/TI, DE) (F) FERMENTATION/TI, DE
 S10 4758 GLUCO/TI, DE
 S11 5239 AMYLASE/TI, DE
 S12 224 GLUCO/TI, DE(W) AMYLASE/TI, DE--- (11)
 S13 1258 (DIGESTI?? *ENZYME/TI, DE + (ETHANOL + ALCOHOL) (F) FERMENTATION/TI, DE + GLUCO() AMYLASE/TI, DE

7c 3*13---(12)

59 3---(13)
 1258 13---(14)

附录: DIALOG 系统部分有关微生物学文献的数据库

文档号	文档名称	主要学科	存储年限
5 55 255	Biosis Previews	生物学文摘	1931—现在 1977—1980 1969—1976
40	Enviroline	环境文摘	1971—现在
34 87 94 186	Scisearch	科学文献索引	1984—现在 1981—1983 1978—1980 1974—1977
76	Life Sciences Collection	生命科学文摘	1978—现在
65	Ssie Current Research	进展中研究项目索引	最近两年
77	Conference Papers Index	会议文献索引	1973—现在
51	Food Science and Technology	食品科技文摘	1969—现在
308 309 320 310 311	CA Search	化学文摘	1967—1971 1972—1976 1977—1979 1980—1981 1982—现在
350 351	Derwent World Patents Index	世界专利索引	1963—1980 1981—现在
233	Telegen	生物工程	1973—现在
285	Biobusiness	生物医学应用文摘	1985—现在
50	CAB Abs.	英联邦农业文摘	1972—现在
10 110	Agricola	农业文摘	1979—现在 1970—1978
60	CRIS/USDA	进展中农业研究项目	最近两年
72 73 172 173	Excerpta Medica	医学文摘	1982—现在 最新 1980—1981 1974—1979

.....

上例中:

- (1) 调用 55 号文档的指令。
- (2) 调用 55 号文档的年、月、日、时间和终端用户号。

(3) 在调用 55 号文档之前用户已占用 DIALOG 系统 1 号文档(通报事项用)的费用和联机机时。

(4) 文档 55 号的名称和存储文献的年限。

(5) 检索系统响应的内容和次序说明: Set 为检索结果的集合号; Items 为文献记录的篇数; Description 为检索提问的标识。

(6) ?ss(Select Steps): 查找文献指令; (t)或 (R) 是位置算符: 表示在此符两侧的关键词必须同时出现在文献记录的同一个字段内, 但两个词的前后顺序不限; ti 或 TI(Title) 是篇名或标题字段的标识符; de 或 DE (Descriptors) 是叙词字段的标识符。标识符是供电脑在检索中识别字段用的符号(下同)。

(7) 检索系统响应的内容, 此后对检索式都给予集合号: S1、S2...

(8) 表示在篇名或叙词中有 RAW 字的文献 1637 篇(下同)。

(9) 输入检索策略指令: “?”是截词 (Truncation) 运算符, 主要是利用某些检索词的词干或不完整词形进行检索, 使所用的检索词包括了该词的词尾的多种变化形式, 从而可减少输入检索词的数量, 节省联机机时, 并可获得较高的查全率。如 digesti?? 可包括 digestive 和 digestion 等。

“*” = AND, 是一种布尔逻辑算符 (Boolean operators), 表示既包括概念 digesti??, 又包括概念 enzy-
me。这样可以缩小检索范围, 减少命中文献。

“+” = OR, 也是一种布尔逻辑算符, 表示“+”

左右两侧检索词中的一个概念或同时包括两个概念, 以扩大检索范围, 增加命中文献。

(10) 同上(7)的解析。

(11) “W”(With) 或“()”是一种位置算符, 表示在它两侧的检索词必须按此前后衔接的顺序排列, 且两词之间不得有其它词或字母, 但允许有空格或标点符号。

(12) ?C (Combine) 为逻辑匹配指令, 用布尔算符对检索词进行逻辑匹配, 形成完整的检索提问, 才能获得所需的检索结果。使用此指令后必须用集合号如 3*13 代替检索词(关键词)。

(13) 集合号 3(S3) 检索提问标识的文献有 59 篇。

(14) S13 的文献有 1258 篇。

(15) 从 55 号文档检索结果, 查到既包含 S3 又包含 S13 检索提问标识的文献 19 篇。

(16) 脱机打印指令。即要求以第五种格式(为二次文献记录的全文, 包括作者、题目、出处、语种和文摘等)打印集合号 S14 命中文献记录的第 1 至第 19 篇。

参 考 文 献

- [1] Dialog Database Catalog, DIALOG Information Retrieval Service, USA, p. 1—4, 1985.
- [2] Dialog “Technical Memo 1—4, Dialog Version 2”, DIALOG Information Services Inc., USA, p. 3—11, 1985.
- [3] System Development Corporation (SDC): Orbit Database Catalog, SDC Search Retrieval Service, USA, p. 7—15, 1982.
- [4] SDC: Orbit User Manual, SDC Information Service, USA, p. 1—7, 1985.