

• 信息管理 •

互联网上生物学信息资源的检索

涂 斌

(陕西省微生物研究所图书馆 710043)

摘要: 本文对因特网上生物学资源的查询途径和利用方法进行了探讨, 介绍了一些较为重要的生物学期刊资源站点, 讨论了如何通过互联网检索各种生物信息资源, 最大限度地利用互联网为生物科学科研服务。

关键词: 互联网; 生物学资源; 检索

Biological Information Resources Through Internet

TU bin

Library of Shaanxi Microbiology Institute, Xian 710043

ABSTRACT: The paper discusses issues on how to retrieval various kinds of biological information resources through internet. Introduces the very important biological resources websites and discusses how to retrieval all kinds of biological information. Hopes to make full utilization of internet to serve and research in bioscience.

Key words: Internet; Biological information resources; Retrieval

在生命科学高速发展的今天, 生物信息成指数增长, 这些信息多数以数据库的形式存储于计算机中, 这些信息资源的优点是方便快捷、易于获取、内容丰富、信息量大。信息查询检索是科研过程中最基本、最常用的步骤之一, 电子文献的易用性和普遍性使人们能够方便地查询到所需要的信息。然而, 正是由于互联网信息资源多, 涵盖的范围广, 使得检索变得很复杂, 检索效率也在不断下降, 互联网的发展趋势是产生越来越多的专业网站, 通过相关专业的网站可以使检索范围大大减少。这里针对我国生物科研中常遇到的问题, 我们对网上众多的生物学搜索引擎进行分析比较, 择其功能强大, 使用方便的部分生物科学网络信息资源以及检索信息的途径和方法进行概括性的介绍。

1 信息检索的途径及需要注意的几个知识点

科学技术高速发展, 目前大部分学科的“半衰期”都在 3—5 年左右, 学科知识的更新速度很快, 通过信息的检索, 我们可以掌握这些新的知识, 那么如何从浩瀚的信息海洋中获取对您有价值的资料呢?

1.1 信息检索的途径一般有六种

①题名途径: 利用书、刊、杂志、文章的名称查找文献, 是最直接、方便的途径。②著者途径: 以作者或团体的名称进行查找。③序号途径: 利用文献的代码、数字编号进行查找。如专利号、科技报告的报告号、技术标准号等。④分类途径: 按文献主题内容所属的学科进行查找。⑤主题途径: 按主题词

的字顺排列进行查找(规范化)。⑥关键词途径: 从文献中选择一定的具有实质意义的词进行查找。⑦其他途径: 如: 化学分子式、地名、属种等途径。

1.2 网络检索中需要注意的几个知识点

①短语检索: 需要用“ ” 来制定检索内容。如“hydatidiform mole”(葡萄胎), 需要用引号标注。②注意大小写的敏感, 输入小写字母, 会同时命中大写字母; 但大写字母却特指大写的单词。因此, 专用名词(人名和地名等)每个单词的首字母必须大写。③检索词的组配和截词符的使用, 检索词前加上“+”表示检出的信息内容中必须包括该单词。例如“+ angioplasty + laser”指检出激光血管形成术的内容; 检索词前加上“-”表示检出的信息必须不包含该单词(切记要在“-”号前留一个空格位)。如“+ hepatitis - animal”指检索出肝炎(非动物类)的内容。注意在使用 Google 查询时不需要使用“AND”, 因为 Google 会在关键词之间自动添加“AND”。Google 提供符合您全部查询条件的网页。“*”为无限截词符(替代多个字母), “?”为有限截词符(替代一个字母), 如“stu*”可以检索出 student, study, studing, stupp 等词。而“stu?”则只能检索出 stud, stun 等词。④逻辑检索 and(与), 如“culture media and agar”指甘油培养基(范围较小)。or(或), 如“culture media or agar”指培养基或甘油(范围较大) and not(非), 如“culture media and not agar”指检出除去甘油以外的培养基。near(临近符)指允许检索的两个单词间最多插入 10 个单词。

2 检索有关生物信息资源网站的导航或门户

(收稿日期: 2006- 02- 06 接受日期: 2006- 03- 12)

关于导航或门户,它并不存贮各类实际的信息资源,它主要是存放有关专业的主题数据库或服务器的地址信息,实际是一种专业学术信息资源的“搜索引擎”,可以指引用户到特定的地址获取所需的信息资源。国内外主要的生物导航或门户有:①生物导航(<http://www.bioguider.com/index.htm>);②国家科学数字图书馆生命科学学科信息门户(<http://www.slas.ac.cn/10000/default.htm>);③生物谷(<http://www.bioon.com/friend.htm>);④生物引擎(<http://www.china991.com/jyhtm>);⑤哈佛大学的生物信息搜索引擎 Biology Links (<http://golgi.harvard.edu>)是互联网生物学的总汇,将生物资源分成:生物学相关机构、在线杂志、模式生物数据库、分子生物学数据库等类,收录全面。

3 利用网上生物信息数据库进行检索

在众多的网站中,“Science”(<http://www.china.sciencemag.org>)可以说是世界上最权威的科技期刊,发表的论文代表当今生命科学的研究方向;(<http://dailynews.muzi.com/tp/chiness/biology.shtml>)“木子新闻”和“Science Daily Magazine”(<http://www.sciencedaily.com/>)以科技新闻为主,从中可以了解世界范围内关于生命科学的最新报道。当然,在因特网漫无边际地漫游中,科技人员深感苦恼的是不知从何处着手找到全文杂志,在这里,我们提供一个 Highwire Press 网站,可以很轻松地找到免费的全文杂志。它是世界上最大的提供免费自然科学文献全文的网站之一,其网址为 <http://highwire.stanford.edu>,主要收集生命科学、医学、物理科学以及少部分社会科学方面的出版物(除期刊外,还有一些非期刊性质网络出版物),由斯坦福大学图书馆 1995 年建立。最初仅出版著名的周刊“Journal of Biological Chemistry”,很快“Science”和“Proceedings of the National Academy of Sciences”也加入其中。收录电子期刊约 282 种,其中 136 种期刊提供免费全文;收录全文 100 多万篇,其中 27 万多篇可免费获得。这些数据仍在不断增加。收录的期刊覆盖的学科有生命科学、医学、物理学、社会科学。检索时可以按学科(View by Subject)、刊名(View Alphabetically)浏览期刊,也可在所有期刊中检索(Search all journals)。其中 Free back issues: 可以免费获得过刊的全文,时间从 3 个月前到两年前,各刊不一。Free trial period: 可以在试用期内获得全文。Free site: 可以获得该刊的所有全文。

4 通过有关专业组织机构网站进行检索

网上与生物相关的国内专业网站主要有:中国医学生物信息网(<http://cmbi.bjmu.edu.cn/>)、中华基因网(<http://www.China.genenet.com/>)、生物虾米服务频道(<http://www.bioxm.com/>)、基因时代(<http://dna99.51.net/>)、中文分子生物学个人交流网(<http://biolover.vicp.net/>)、生物通(<http://www.ebio-trade.com/>)、生命释站(<http://netroom.hbu.edu.cn/personal/gene2000/>)。国内外的一些生物信息中心是生物科学专业资源的集聚地。较重要的国外的生物科学的组织机构站点主要有:NCBI,美国国家生物技术信息中心(www.ncbi.nlm.nih.gov),所有公开的 DNA 序列的集合,世界上的权威序列数据

库,EMBL,欧洲分子生物学实验室([WWW.embl-heidelberg.de/Services](http://www.embl-heidelberg.de/Services))、EBI,欧洲生物信息学研究所(<http://www.ebi.ac.uk>)、NCGR,美国国家基因组资源中心(www.ncgr.org/)、ICGEB,国际遗传工程与生物技术中心(www.icgeb.trieste.it)、林奈的生物信息中心(www.se.embnet.org)。国内的生物信息中心起步较晚,目前主要有 CBI,北京大学生物信息中心(ICP, <http://www.lpc.pku.edu.cn>)、BIOSINO,中国科学院上海生命科学院生物信息中心(www.biosino.org),此外,一些图书馆和文献中心也是重要的生物信息资源的收藏机构,如:全国医学文献信息中心:北京医科大学图书馆(<http://library.bjmu.edu.cn>),国家科技图书文献中心(<http://www.nstl.gov.cn/index.html>)等。

5 生物信息的检索工具

5.1 专业的检索系统

①Entrez。美国国家生物技术中心 NCBI 的 Entrez 是目前国际上最为著名的生物信息数据库查询系统,用于对文摘摘要、序列、结构和基因组等数据库进行关键词查询,找出一个或相关的数据库条目(www.ncbi.nlm.nih.gov/Entrez/)。

②SRS。SRS 是由欧洲分子生物学实验室开发,目前是欧洲各国主要生物信息中心必备的数据库查询系统。SRS 是一个开放的数据库查询系统,即不同的 SRS 查询系统可以根据需要安装不同的数据库,目前共有 300 多个数据库安装在世界各地的 SRS 服务器上。

5.2 生物信息分析软件

网上有大量生物信息专业软件,还有一些免费软件的档案库,BLAST 和 FASTA 是使用得最为频繁的两套数据库搜索程序,一般认为,BLAST 软件是较通用的一种同源性检索工具,由一种功能不同的程序组成,运行速度快,对蛋白质的搜寻更有效;FASTA 运行较慢,对核酸序列更为敏感,所以说,只要条件允许,就应当 BLAST 和 FASTA 双管齐下。另外,网上还有其他一些关于生物信息软件的网站,值得一提的是中国的一个网站—生物软件网(www.bio-soft.net)它是华北制药集团的谈杰创建的一个网站。随着人类基因组计划的开展,争夺有限基因资源的竞争日益激烈,利用网上资源分析未知序列,迅速得到尽可能多的信息尤为重要。现在一些这方面的软件(<http://bioinformatics.weizmann.ac.il/gdp/yd.html>)根据功能归类被集中在一个网页上,极大的方便了使用。

参考文献

- [1] 刘同奎. 医学图书馆的网络信息资源与利用. 生物磁学, 2005, 5 (3): 93- 95
- [2] 安俐颖. 建立医院图书馆电子阅览室的体会. 生物磁学, 2005, 5 (4): 90- 91
- [3] 杨世成, 刘秀英. 国际互联网上生命科学信息资源. 生物学通报, 2003, 38(1): 21- 22
- [4] 李建波. Internet 上的生物信息资源及获取. 生物学通报, 2002, 37 (7): 30- 31
- [5] 陆丽娜, 郭金丽, 于啸. 网络环境下生物信息资源的检索. 农业图书情报学刊, 2004, 15(6): 74- 75
- [6] 杨凝青. 因特网生物医学期刊资源搜索与利用. 图书馆杂志, 2002, 21(12): 56- 60, 63