

互联网上的免费 学术信息源及其获取

凌美秀 曹春晖 主编



湖南大学出版社

互联网上的免费学术 信息源及其获取

主编 凌美秀 曹春晖

湖南大学出版社

2007年·长沙

内 容 简 介

本书提供的网络免费学术信息源,包括各种全文信息源、网络数据库以及能提供信息检索的检索工具与检索系统,涉及字/词典、百科全书、年鉴、手册、指南、年表、历表在内的许多免费参考信息源。可作为从事科学研究、科技创新、教学等工作的读者获取网上免费信息资源的操作指南。

图书在版编目 (CIP) 数据

互联网上的免费学术信息源及其获取/凌美秀,曹春晖主编.
—长沙:湖南大学出版社,2007.3

ISBN 978-7-81113-170-3

I. 互... II. ①凌... ②曹... III. 互联网—情报检索

IV. G252.7

中国版本图书馆 CIP 数据核字 (2007) 第 026245 号

互联网上的免费学术信息源及其获取

Hulianwang Shang de Mianfei Xueshu Xinxiyuan Jiqi Huoqu

作 者: 凌美秀 曹春晖 主编

责任编辑: 罗素蓉

封面设计: 张 毅

责任校对: 祝世英

出版发行: 湖南大学出版社

社 址: 湖南·长沙·岳麓山

邮 编: 410082

电 话: 0731-8821691(发行部), 8823113(编辑室), 8821006(出版部)

传 真: 0731-8649312(发行部), 8822264(总编室)

电子邮箱: pressluosr@hnu.cn

网 址: <http://press.hnu.cn>

印 装: 长沙瑞和印务有限公司

开本: 850×1168 32 开

印张: 7.25

字数: 176 千

版次: 2007 年 3 月第 1 版

印次: 2007 年 3 月第 1 次印刷

书号: ISBN 978-7-81113-170-3/G·289

定价: 16.00 元

版权所有, 盗版必究

湖南大学出版社凡有印装差错, 请与发行部联系

前 言

一直以来,“天下没有免费的午餐”不只是成了人们不相信各种免费服务的口头禅,也是大多数人进行价值判断的通行依据。随着互联网进入并深深融入现代人的生活,她也为结束人们这一思维定势带来了契机。当你带着耳机一边轻松地欣赏从网上免费下载的流行歌曲,一边享受搜索引擎提供的网络信息免费搜索服务之时,你一定找不到反驳这一结论的任何理由。如果你从互联网上发现一本一直想购买图书的电子全文,或者是找到一组极具参考价值的学术文献,千万别以为这是自己的好运气使然。要知道,免费的学术信息源已经是网络信息源重要的、不容被忽视的组成部分。

每一个从事学术研究、科学创新的用户,在进行自助式学术信息资源检索时,往往将最先的选择锁定在图书馆入藏的印本和数字信息源。由于知识产权保护方面的原因,这些资源基本上只限于向特定局域网上的用户提供。而包括电子图书、电子期刊、会议论文、学位论文、专利与标准信息免费信息源,各种免费的网络数据库,由电子词典、百科全书、年表与历表、手册与指南构成的无处不在的免费参考信息源与参考工具书,则是没有任何限制地向全体网络用户免费开放,并且在数量与品种上呈递增趋势。这些提供免费学术信息源的机构与个人不只是缩小信息差距,维护信息公平的倡导者,更是积极的实践者,他们是现代社

会最值得尊敬的人。

至于本书和它名不见经传的作者，充其量只是充当了资源发现者的角色。当本书从一大堆零散的资源中化蛹成蝶之时，笔者依然能清晰地记起每发现一个免费的资源网站所带来的窃喜。当然这种发现还是基于国内的同行们付出的前期努力。正是通过国内已经建成的学科导航库与学科信息门户，通过本书参考引文中列出的诸多学者的无声的指引，本书的作者才有可能将这部有关互联网上免费学术信息源的专著奉献给各位尊敬的读者。遗憾的是，期望将网上的免费资源一网打尽，只能是一种不切实际的幻想，尽管这曾是本书作者最初的愿望。

最后，作者要感谢郑章飞教授在本书定题与成书的过程中给予的鼓励与指点，感谢同事李雯老师的无私帮助。由于水平所限，本书的错误与不足之处在所难免，恳请专家学者提出改进意见，我们将十分感激！

凌美秀 曹春晖

2006 年冬于岳麓山

目 次

第一章 网络免费学术资源概述	1
第一节 网络免费信息源概况/1	
第二节 互联网上的免费学术资源/5	
第三节 网络免费学术资源的类型/10	
第四节 网络免费学术资源的来源/18	
第二章 互联网上的免费全文信息源	23
第一节 免费全文图书/23	
第二节 免费电子全文期刊/33	
第三节 免费网络报纸/59	
第四节 免费学位论文/61	
第五节 免费专利信息源/62	
第三章 免费网络数据库	66
第一节 用于图书检索的免费数据库/66	
第二节 用于期刊检索的免费数据库/69	
第三节 用于学位论文检索的免费数据库与检索系统/71	
第四节 用于会议论文检索的免费数据库与检索系统/73	

- 第五节 用于专利检索的免费数据库与检索系统/77
- 第六节 用于标准及相关信息检索的免费数据库/83
- 第七节 重要的免费专题特色数据库/94
- 第八节 其他免费的网络数据库/125

第四章 免费参考工具与参考信息源

第一节	网络免费参考工具与参考信息源概况/136	136
第二节	网络参考工具书的获取途径/137	
第三节	免费字、词典/140	
第四节	免费百科全书/150	
第五节	免费年鉴/158	
第六节	免费年表与历表/166	
第七节	免费手册与指南型参考工具/168	

第五章 免费网络信息资源检索工具与检索平台

第一节	网上免费论文全文检索系统——预印本系统/170	170
第二节	Google Scholar ——专门提供网上学术资源搜索服务的工具/174	
第三节	开放存取——完全免费的开放式学术信息资源共享计划/178	
第四节	网络信息检索工具——搜索引擎 (Searching Engines) /184	
第五节	网络信息资源重点学科导航库/201	
第六节	学科信息门户/210	

参考文献

	224
--	-----



网络免费学术资源概述

第一节 网络免费信息源概况

互联网上提供信息内容服务的主体具有多样化的特征, 各种以网络为平台发布、传输信息的机构与个人都是网络信息的提供者。在众多的信息资源与内容提供商中, 有数据库生产商、学术期刊出版商、各种行业性的学会与协会、商业性的联机检索系统(如 Dialog、STN), 以网络为平台提供中介性信息服务的数字图书馆, 以及进行网络信息的采集并提供网络信息连接服务的搜索引擎。网上学术信息的种类多样、分布广泛。这些专门从事网络信息服务的机构与组织, 占据了互联网上学术信息交流与服务的主体。但是, 数据库生产商、电子文献的出版商出于自身利益的考虑, 它们提供的电子文献的全文信息是收费的, 能够对全体网络用户免费开放的仅限于题录与摘要信息。而以图书馆为代表的公益性信息服务机构, 因受制于商业性电子资源全文出版商的IP限制, 所服务的对象有着极为明确的权限差异, 绝大部分网络用户不能取得对商品化数据库的合法使用权, 能提供免费服务的仅限于联机公共目录查询、学科导航库、学科信息门户、在线数字参考咨询等。通常情况下, 互联网上的资源大多数是收费的, 尤其是高品质的学术资源。提供免费资源的网站虽然对其用户不收费, 但它的存在需要有资金支持。从提供免费资源的主体

类型看,可以把所有提供免费资源的网站分为两大类,一是有资金支持的网站,如政府部门的网站、企业的网站、教育机构的网站和科研机构的网站等等,一般来讲其内容都是为提供支持资金的机构服务的;另一种是没有资金支持的网站,这些网站要存在,就要有一个盈利的模式。对于向全体网络用户提供免费资源与服务的网站而言,广告收入是其盈利构成中的主要来源。要得到更多的广告收入,其目标就是要获得更高的点击率,这样的盈利模式就决定了这些需要自己盈利的网站提供的内容一定是大众化的、普及性的。因此,提供学术性免费资源的信息服务商在数量上要远远低于通俗类免费信息提供商。

一、盈利性网站的免费信息源

提供免费资源的盈利性网站盈利来源主要靠广告收入。通过提供免费信息获得广泛的影响,提高网站内容的点击率,锁定网络用户通常是这类网站的通行经营模式。根据这个目标,网站把所有的上网者都作为其潜在的访问者,所以提供的信息和资源具有大众化的特点。虽然从网站数量上来讲,盈利性网站并没有占网站的多数,但从点击率上讲,盈利性网站却占了大多数,是上网者关注最多的网站。

1. 免费新闻信息源

新闻和综合信息网站是网上免费资源中访问量最大的网站。因为这些内容是最受关注的信息之一,适用于每一个上网者。提供新闻和综合信息的网站与印刷型的报纸及其他媒体具有相通之处。从内容传播的速度上,网络比印刷型报纸有着更多的优势;从容量上讲,新闻和综合信息网站突破了报纸等传统媒体在版面上的容量限制。但在阅读成本上,网站处于明显的劣势,观看网

站所提供的新闻需要一台联网的计算机,其费用包括计算机的费用和上网的费用,与一般为一元以内的报纸的价格相比,其费用是非常高的。从阅读的便利上讲,新闻和综合信息网站也不像印刷型的报纸那样可随身携带。因此新闻网站虽然从内容、速度和容量上来讲都相似于或超过印刷型的报纸,但在阅读成本和方便程度上不如报纸,所以新闻网站并不能代替印刷型报纸。现在各大报纸也都开设了自己的网站,由于与专业新闻网站有不同的盈利模式,这些网站大都没有铺天盖地的广告,更加使专业新闻性网站在竞争中处于劣势。新闻信息类包括两大内容板块:新闻与综合信息。大多数网站将新闻信息其内容区分为社会、财经、体育、娱乐、科技等方面。综合信息包括一些热门的内容,比如教育、汽车、住房、健康、旅游、电脑等,这些虽然具有一定的新闻性,但是较多的是产品及机构的介绍和评论,所以不能简单地称之为新闻。

2. 文娛休闲类免费信息源

文娛休闲类网站以文学站点和论坛形式的网站为代表。文学站点的主要内容有古今文学名著、时尚小说、网络文学三部分,其中的核心是网络小说。网络文学的影响力已经达到为正式出版机构所接受的程度,网络文学的质量已经达到了与其它方式创作的文学作品相当的水平,以至于现在已经有出版社将网上原创的小说以“网络小说”为丛书名结集出版。论坛的交互式方式是网站本身的特点形成的一种文娛方式。在同一个主题下,让每个人根据自己所掌握的知识,发表自己独特的观点,这种形式也吸引了非常多的人参与到这个内容中来。但由于网上发表的见解没有一个质量上的审核过程,因而在信息分类体系中,把论坛作为文娛类而不是普及类。文娛类的网站虽然面对的对象也是所有的上

网者，就其内容来说，也适用于所有的人。但娱乐方式是多种多样的，所以只会有一部分的人选择这些网站作为娱乐，因此其影响力远远不如新闻类的网站。

3. 通俗普及型免费信息源

提供各个学科基础普及知识的网站，主要包括知识性和学习性两类。这些网站往往选取一个主题作为网站的内容，比较热门的有军事、法律、电脑、艺术等。该类网站所面对的对象是所有的上网者，但由于学科范围及用户专业背景的差异，影响力与点击率低于前几类网站。

二、非盈利性网站的免费资源

非盈利性网站是互联网中数量最多的网站，包括各种基金会、学会、协会的网站。这类网站只能锁定部分读者，因此其点击率远远比不上盈利性网站。主要原因是非盈利性的网站是为向其提供资金支持的机构服务的，其发布的内容往往具有专指性，只有在需要查找相关信息时，上网者才会使用这些网站，所以这些网站并不为大多数上网者所经常关注。非盈利网站提供的信息源主要有以下几类。

1. 机构企业信息类

由于互联网广泛的传播性，各种机构，包括政府机构、教育机构、文化机构等等都在网上设立了介绍本机构和发布本机构信息的网站。这些机构网站的加入极大地丰富了互联网的资源，方便了人们更快捷地获得各种信息。

2. 学术类

学术类资源在互联网上的比例并不高,但由于互联网容量巨大,使得互联网上学术资源的绝对数量并不少。因此,学术资源的检索与获取难度比较大,数据库的建设成为管理与收集学术信息源的重要方式。这一特征为各类数据库生产商提供了无限的商机。网上学术类资源的提供机构各异,专业化数据库生产商提供的学术资源大多是收费的。同时也提供部分免费信息作为开发市场的一种手段。收费信息网站中提供的免费信息,主要来源于研究机构、图书馆、学术期刊网站等。

第二节 互联网上的免费学术资源

一、网上免费学术资源的界定

免费的网络学术信息源是指那些不需要网络用户支付任何费用,并且无限制地面向全体网络用户提供信息内容获取服务的所有类型资源的总称。提供免费的学术信息资源的信息内容提供商,网络用户对所提供信息资源的在线浏览与下载等一切利用方式的一切信息内容服务提供技术支持。从网络免费信息源提供方是否通过信息内容提供服务取得商业性收益的角度,可以将提供信息内容获取服务的网站分为盈利网站与非盈利网站两大类。从资源的属性看,免费资源可以分为学术资源与非学术资源两大类。从资源锁定的用户群体看,学术型资源主要面向专家、学者、专业研究人员、教师及具有特定学科背景,且对资源的选择具有明显的学科倾向等网络信息用户。通常,盈利型的网站以获得更高点击率为目的,提供的内容主要是适合大众的新闻、热点

及相关信息、文娱性的以及普及性的知识等。根据其内容类型的不同,可以区分为三类,新闻信息类、文娱类和普及类。第二大类是有资金支持的非盈利性的网站。其一是各种机构和各种企业介绍本机构并发表信息的网站;其二是发布学术性内容的网站。

网上丰富的学术信息资源涵盖了各个学科专题与各类按出版类型进行区分的各种文献类型。在学术信息源中包含了大量通过其他媒体很难获得的信息资源,如行政公文、统计数据、行业标准、产品说明等。依据网络共享学术资源集群和资源类型,目前网上的学术信息基本上可分为:新闻信息、动态信息、规范出版的全文信息和书目信息、数据库、参考信息源六大类。新闻信息中具有学术价值的是各种专门性电子媒体(门户网站、与教学科研有关的机构网站)、传统新闻媒介的网络版(如人民日报网络版)发布的各种新闻中对学术研究有参考价值的部分信息。这部分信息对科学工作者了解科学发展动态、科学研究政策、社会与经济发展对科学研究的影响等具有不可忽视的作用。网上动态信息是充斥于网上最广泛的信息资源,其中具有学术价值的主要是一些学术机构发布的有关学术活动信息、学术研究个人的某些可以公开的信息、企业有关产品研发和技术支持信息等。这两类信息基本上都属于简单报道和发布,信息的产生缺乏有效的监督和权威的审核与评价,查询行为随机性和盲目性大,无法判别查询的全面程度如何。因此,网上学术信息的主要来源为规范出版的全文信息和书目数据库,它包括图书信息、期刊论文信息、会议信息、专利标准信息、学位论文信息、科技报告信息、政府出版物信息等。网络学术资源按出版正式程度可分为非正式出版的信息,半正式出版的信息,正式出版物三大类。非正式出版的信息包括电子信函、个人主页上表述个人观点和见解的非正式出版论文、学术论坛上的文章和其他信息;半正式出版的信息包括政

府工作报告、机构工作进展报告、统计数据、教学大纲、产品说明、样品报道、会议通报,大都可以从政府机构、国际组织、学术团体、教育机构、企业商业部门等的网站上免费获取。正式出版物包括:电子期刊(纯电子期刊、纸本期刊的网络版)、数据库、电子图书等。

二、网上免费学术资源的特点

网上免费学术资源的特点,可以通过与图书馆数字资源的比较得出。后者的最显著的特征在于有一个具有明显标识的信息门户,这个门户便是特定图书馆的网站。这一特点保证了以特定数字图书馆为平台所聚集的学术信息源的集中度,为用户提供了统一的资源获取门径。而分散于整个网络、以异构性为特征的免费学术资源则呈极度的分散分布状况。海量资源以有序化为利用的前提,网络资源的有序化是通过合乎用户利用习惯的资源组织来完成的。对网上免费学术资源的检索主要是通过专门提供网络信息检索的网络信息搜索工具。YAHOO, GOOGLE 等较著名的专业检索网站均属此类,以这些专业检索网站为代表的互联网检索在查全率上的表现是比较好的,但是在查准率方面表现得并不尽如人意。造成这种现象最为主要的原因是互联网上学术资源占的比重很小。总的来说,网络免费信息资源的特征可以用“分布分散,品种多样,获取困难”来描述。

网络免费信息源来源广泛,造成了其分布的无序与分散。政府机构、数字图书馆、研究机构、教育机构都是免费资源的发布者。除网络图书、网络期刊、会议论文、学位论文、专利信息、统计数据外,还有论坛、个人主页、博客等。可以说,在种类上网络免费信息源几乎涵盖了一切资源门类。到目前为止,还没有一个专门提供免费学术信息源检索的资源门户。而且,提供免费

学术资源的网站也不仅仅以提供该类资源为特征。这一系列的因素造成了免费学术资源的极度分散。同时,由于免费学术资源本身表现在种类上的多样性,造成网上免费学术资源的获取困难,以至于通过通用性的综合搜索引擎很难获得满意的检索效果。以GOOGLE为例,输入一个关键词,其检索结果可能是上万条。除了检索结果在数量上多到用户无从筛选,另一个常常被网络用户抱怨的原因是,依靠搜索引擎作为网络信息源的采集工具往往造成对检索结果的质量判断缺少依据,及对信息等级无法作出相应的判断。检索系统本身除提供检索结果的排序方式选择外,无法判断资源与用户检索需求的内容关联度。造成了网上学术资源获取必须通过人工筛选,这是网上免费学术资源的一个主要特点。提高网上免费学术资源的查准率与查全率的唯一措施,是对所收集的信息源进行人工介入的规范化处理。因此,提高免费信息资源的查准率与查全率,一直以来便成为专业信息搜索服务提供者的努力方向。以数据库为主的图书馆提供的数字资源检索设计了多途径的检索入口。一般来讲,数据库中的资源都有较完整信息组织方案以及行之有效的信息检索系统,成为到目前为止收集与整合免费学术资源的最好的工具。

与庞大的网络信息源相比,网络免费学术资源在总体数量显得极为有限。因此,对学术研究而言,总体上它只起到补充的作用。这一结论并不适用于所有学科与研究领域,对于某些研究领域,网上免费学术资源的作用不可轻视。如医学信息资源,在网络免费信息源中所占的比例相对而言较之其他学科门类要大得多,研究者对一些学科的免费网上资源作了挖掘,从中发现一个很大的特点,90%以上的文献对医学和相关学科,如生物医学、生命科学等。这种现象说明医学是网上免费资源最多的学科。Free medical journals 网站认为网上免费的医学信息多于其

他研究领域的原因在于：免费获取科学文献对医学实践有很大的作用，医学实践不仅仅是一种研究，也是一种救死扶伤的公益事业，这可能就是很多的网站致力于收集、整合网上免费资源提供给医学工作者利用的主要原因。Free medical journals 网站现已收集了 970 种网上免费医学资源，并认为在未来的几年里，人们将能从网上免费获取更多重要的医学期刊全文。众多的文献提供了多个获得医学文献的途径，比如美国国家医学图书馆的网站、美国医学学会的网站、英国医学杂志出版集团的网站等等。这一特点也反映在对不同学科领域的网络免费资源的挖掘程度上。鉴于这些情况，图书馆领域的研究者对网上免费学术资源进行了多种形式的组织与挖掘工作。

在免费网络学术信息源的数据类型看，有题录型、文摘型、全文型信息源。其中，题录与文摘类免费信息在数量上远远超过全文型信息源。目前提供学术期刊免费题录的网站很多，一些大型的全文数据库都提供免费信息供使用者了解自己的数据库，这使得取得免费的题录和文摘是比较容易的事。但是提供免费全文的网站相对比较少，一般提供全文有两种情况，一是提供非即时信息，这是比较多的情况；另一种是提供即时免费全文的情况，这种情况较少。这些全文往往是各个学术期刊在网上提供的电子版，信息比较零散，需要人工的方式收集，然后提供给使用者使用。也有网站专门从事整合这些信息的工作，特别是如前所述的医学学科，有较多的收集免费资源提供给研究者使用的网站。

从免费网络信息源涉及的出版类型看，则有图书、期刊论文、学位论文、专利信息、标准信息、统计信息等。图书与网络期刊在数量上占有压倒一切的优势，这一点，与这两种文献类型在传统文献群体中所处的地位是高度一致的。

第三节 网络免费学术资源的类型

如前文所述, 互联网上的免费学术资源有书目信息、网络数据库、电子图书与电子期刊、网络报纸、网络参考资源、软件资源、动态性信息(如BBS、新闻)、科技报告、会议文献、学位论文、专利、技术标准、政策法规、统计资源、行业报告等。这些免费学术信息资源以多媒体、超链接、超文本为存在特征, 且来源与出版类型各异。从提供信息的服务器类型看, 有在构成比例上占目前网络信息资源绝对优势的Web信息, 有由FTP(Transfer Protocol 文件传输协议)文件服务器提供的FTP资源及其他资源; 从出版类型上分, 有电子图书信息、电子期刊与期刊论文信息、会议信息与会议论文、以网络工具书为主体的参考信息源、专利信息源、科技成果与科技报告、产品信息、档案等。此外, 还有以动态信息为主体的非结构化信息, 依托各种综合性与专业性网站存在的新闻信息; 从功能与加工层次看, 既有图书、论文等全文信息, 更多的则是提供全文获取线索的题录、目次、文摘等书目信息, 在三次文献的家族中, 有在线字/词典、网络百科全书、网络年鉴、名录信息、手册等网络工具书, 这些网络信息也被称为参考信息源。

一、网络图书与图书信息

互联网上免费的图书信息主要为书目信息, 有题录型书目信息与文摘型书目信息之分。书目信息之外, 也有部分的全文信息。书目类图书信息主要由图书馆、出版社(商)、图书网络发行商提供。图书馆提供的是馆藏印本图书的联机公共目录查询, 图书出版商/出版社往往通过自己的网站发布所出版图书的书目