

Internet 信息资源检索丛书

化学化工信息资源检索和利用

袁中直 肖 信 陈学艺 编著

江苏科学技术出版社

图书在版编目(CIP)数据

化学化工信息资源检索和利用 / 袁中植等编著.
—南京: 江苏科学技术出版社, 2001.5

(Internet 信息资源检索丛书)

ISBN 7-5345-3353-8

I. 化... II. 袁... III. ①因特网-化学-情报检索
②因特网-化学工业-情报检索 IV. G252.7

中国版本图书馆 CIP 数据核字(2001)第 028125 号

Internet 信息资源检索丛书

化学化工信息资源检索和利用

编 著 袁中直 肖 信 陈学艺
责任编辑 龚福亚 王明辉

出版发行 江苏科学技术出版社
(南京市湖南路 47 号, 邮编: 210009)

经 销 江苏省新华书店
照 排 南京展望照排印刷有限公司
印 刷 淮阴新华印刷厂

开 本 787mm × 1092mm 1/16
印 张 20
字 数 500 000
版 次 2001 年 5 月第 1 版
印 次 2001 年 5 月第 1 次印刷
印 数 1—4 000 册

标准书号 ISBN 7-5345-3353-8/TQ·30
定 价 26.00 元

我社图书如有印装质量问题, 可随时向我社出版科调换。

前言

Internet 是由分布于全球 员多个国家和地区的无数计算机组成的一个巨大网络,在其海量的信息中,化学化工及其相关信息资源就犹如大海中的小鱼,分布杂乱,变化快,极具动态性,难以捕捉,在出海捕获这些资源之前,你必须先清楚以下三个问题:

员 What? 找什么? 网络上的化学化工信息资源有哪些? 什么是你感兴趣的内容?

圆 How? 如何找? 如何有效地找到你所需要的信息资源?

猿 Where? 从哪里找? 为你提供信息资源检索服务的工具在哪里?

Internet 最大的特点就是资源共享,谁能有效地利用这一宝藏,就看谁发掘 Internet 资源的手段快速准确。本书就是从上述三个方面专门论述 Internet 化学化工信息资源的检索与发掘利用,使读者真正达到‘秀才不出门,能知天下事’的境界。我们相信,Internet 将从根本上改变我们检索文献资源的习惯!

本书是一本教材。所有文献、资源、检索工具都采用‘解剖麻雀’的方法,选择 员个典型进行重点介绍和剖析,结合化学检索课题实例,举一反三地使读者学会这一类资源的检索方法。每一节都在最前面列出了本节重点,使读者更快、更容易地了解该章节主题,同时在每节的最后还列举了练习题,读者可以用作上机实习的课题。

本书是一本手册。其中不但有典型资源的使用方法介绍,还包含了大量的 Internet 资源的 URL,这些资源都经过作者精心评价和筛选,指明了这些资源的特点及其与化学化工专业的相关性,便于读者从中发掘更多、更广阔的,适合自己研究特点的 Internet 资源及其检索方法。

本书是一本研究 Internet 资源检索方法和技术的研究文集。全书各章都具有研究性质,集合了 Internet 资源的发展状况、Internet 资源的收集、标引和检索技术,以及在 Internet 资源中检索专业化的化学化工资源的特有方法和提取技术等方面的研究状况和前沿领域。其中也包含了本书作者的研究成果。

本书是动态资源的指南针。Internet 资源极具动态性,随时都在发生变化,因此本书的重点是讲述 Internet 资源的检索方法,并不是 Internet 资源的罗列,更不是网址大全,是一本在一定时期内都不会过时的指南性的书,这也是我国第一本系统地论述 Internet 化学化工信息资源检索方法的书。为了保证这些方法的时效性和最大限度地为读者提供交流、咨询的方便,我们专门为此书设立了两个 WWW 网络支持站点,“化学世界”<http://www.chemworld.com.cn/>和“化学之门”<http://www.chemdoor.com/>,读者可以从这两个站点上获得免费的检索技术支持和追踪本书资源的动态变化,因此本书又是一本动态的书,是一本有 Internet 网络支持的书。

Internet 资源检索是一门实践性很强的学习课程,只看书是无法理解其使用方法和技巧的,只有边看书边实践,才能有针对性地学习,学有所用,熟能生巧。

本书的写作就是一个网络资源利用的好故事。本书的几位作者以及本书的编辑、出版者至今未见过面,因此我们只能称为网友。我们都是较早‘触网’的化学化工专业工作者,我们也都是在检索 Internet 化学化工资源时发现了对方建立和维护的化学化工专业站点,并互相仰慕已久,就是未曾谋面。直到有一天,出版社希望能够出版一系列的专题性 Internet 信息资源检索的书籍,发 Email 给我们这几位 Webmaster 约稿,我们才一拍即合,从此通过‘伊妹儿’信来信去,互相讨论和切磋,最终成就本书。整个过程都是无纸化、电子化或数字化、网络化的写作与协作,因此书中的有关 URL 网址等都是正确的和经过验证的,但在文字写作、逻辑条理上难免有不少错误之处,在重要资源的采集和筛选方面也有偏颇和遗漏,请读者毫不犹豫地给予批评指正,直接发 Email 给作者(yuanzhongzhi 岳china.com),或在前述网站上获取最新信息。

作 者

圆 圆 年 月 于广州

目 录

员	Internet 资源的构成	员
员	Internet 资源的构成及其信息量分布	员
员	Internet 资源定位器——URL	猿
猿	检索 Internet 化学化工信息资源所需的软件	源
圆	WWW 资源	远
圆	什么是 WWW?	远
圆	WWW 域名与化学信息的关联	远
猿	FTP 资源	怨
猿	什么是 FTP?	怨
猿	FTP 资源的搜索	园
猿	常用的 FTP 工具和文件下载工具	园
源	基于用户交流的信息资源	猿
源	Email 及其应用软件	猿
源	BBS 电子公告板系统	源
猿	Usenet/Newsgroups 信息系统	猿
源	Listserv/Mailing List 信息资源	园
源	订阅化学化工专业站点的电子邮件刊物	园
源	Internet 信息交流常用工具软件(BBS/Email/News)	园
缘	Gopher/Telnet/WAIS 资源	猿
缘	Gopher、Telnet、WAIS 资源	猿
缘	Gopher、Telnet、WAIS 资源的检索工具	猿
园	化学化工信息资源的一般检索	猿
圆	Internet 资源检索方法概述	猿
圆	Internet 化学化工信息资源的分类	猿
圆	Internet 资源检索系统	猿
圆	Internet 资源检索方法和检索技术	猿
圆	通用型搜索引擎	源
圆	搜索引擎典型介绍——AltaVista	源
圆	几个重要搜索引擎的比较	源
圆	其他搜索引擎	源
圆	主要的中文检索工具简介	缘
圆	搜索引擎的跟踪	缘
猿	通用型分类目录索引	缘
猿	Yahoo!	缘

猿猿圆	Galaxy	猿缘
猿猿猿	eBLAST	猿苑
猿猿源	其他通用型分类目录索引	猿愿
猿猿缘	自然科学类分类索引工具	猿愿
猿源	集成式检索工具	猿远
猿源员	检索工具的集成方法	猿远
猿源源	典型的集成式检索工具 MetaCrawler 简介	猿源
猿源源	其他集成式检索工具	猿缘
猿源源	桌面上的集成式检索工具	猿远
猿缘	化学化工资源的综合性导航系统	猿愿
猿缘员	Internet 化学化工专业化资源检索工具的现状	猿愿
猿缘圆	典型的化学化工资源导航站点介绍	猿愿
猿缘猿	其他著名的化学化工资源导航系统	猿缘
猿缘源	中文界面的化学化工资源导航站点	猿苑
猿远	化学化工虚拟社区(一)	猿愿
猿远员	化学化工虚拟社区的概念	猿愿
猿远圆	ChemWeb.com	猿愿
猿远猿	美国化学会与 ChemCenter	猿愿
猿苑	化学化工虚拟社区(二)	猿愿
猿苑员	英国皇家化学会与 ChemSoc	猿愿
猿苑圆	德国的 Chemie.de	猿愿
猿苑猿	CS ChemClub	猿愿
猿猿	化学化工资源分类检索	猿愿
猿猿员	Internet 化学化工专利资源检索	猿愿
猿猿员员	专利检索步骤	猿愿
猿猿员圆	IBM 知识产权网 IPN	猿愿
猿猿员猿	美国专利商标局的 WWW 专利服务	猿愿
猿猿员源	QPAT-US 的专利摘要库	猿愿
猿猿员缘	欧洲专利局的 esp@cenet	猿愿
猿猿员远	加拿大专利数据库	猿愿
猿猿员苑	Derwent 的 WWW 专利服务	猿愿
猿猿员愿	中国专利检索数据库	猿愿
猿猿员怨	专利资源检索的其他途径	猿愿
猿猿圆	Internet 上的期刊资源	猿愿
猿猿圆员	Internet 上化学期刊的形式	猿愿
猿猿圆圆	科技期刊文献的检索工具	猿愿
猿猿圆猿	中国的科技期刊检索	猿愿
猿猿猿	化学化工标准及其他文献资源	猿愿
猿猿猿员	Internet 上的化学化工标准资源	猿愿
猿猿猿圆	政府报告资源及其检索	猿愿
猿猿猿猿	学位论文资源及其检索	猿愿

猿猿源 会议文献资源及其检索	猿猿
猿源 化学化工软件资源	猿源
猿源员 计算机软件在化学化工领域的应用	猿源
猿源圆 软件资源搜索	猿源
猿源猿 化学软件导航站点	猿猿
猿源源 著名的化学软件开发公司	猿源
猿源缘 软件编程经验与技巧的交流与讨论	猿源
猿缘 化学教育资源	猿缘
猿缘员 化学教育资源导航系统	猿圆
猿缘圆 化学远程教育资源	猿远
猿缘猿 其他化学教育资源	猿源
猿远 化学物质结构资源(一)	猿远
猿远员 物质结构与命名	猿远
猿远圆 物质结构的绘制与建模软件	猿远
猿远猿 物质结构的可视化	猿源
猿远源 物质结构信息在网络上的传输	猿源
猿远缘 物质结构数据库	猿远
猿苑 化学物质结构资源(二)	猿远
猿苑员 物质结构的光谱分析	猿远
猿苑圆 物质结构的解析与预测	猿员
猿苑猿 基于物质结构的文献检索	猿猿
猿愿 化学参考工具资源	猿愿
猿愿员 NIST Chemistry WebBook	猿远
猿愿圆 Beilstein Information Systems	猿员
猿愿猿 在线工具书	猿圆
猿愿源 化学反应数据库	猿猿
猿愿缘 物性数据和工程数据查询	猿猿
猿怨 化学工商业信息资源	猿怨
猿怨员 化工综合性搜索引擎	猿怨
猿怨圆 化学工商信息数据库	猿猿
猿怨猿 网络化工商务	猿远
源 基于 Internet 的联机文献检索	猿员
源员 数字图书馆资源	猿员
源员员 数字图书馆概念及发展概况	猿员
源员圆 图书馆藏资源的在线查询方法	猿猿
源员猿 OCLC 联机计算机图书馆中心	猿源
源圆 基于 Internet 的联机文献检索系统	猿圆
源圆员 STN 国际联机检索系统	猿源
源圆圆 DIALOG 国际联机数据库检索系统	猿员
源圆猿 EBSCOhost 系统	猿猿
源圆源 其他基于 Internet 的联机检索系统	猿源

源猿	传统化学文献检索工具的在线查询	猿源
源猿员	美国《化学文摘》(CA)及化学文摘社(CAS)	猿源
源猿圆	美国《工程索引》(EI)	猿缘
源猿猿	美国 ISI 的数据库检索	猿源
缘	网络资源检索的辅助技术与前景展望	猿缘
缘员	资源检索中的语言解决方案	猿缘
缘员员	桌面软件解决方案	猿远
缘员圆	在线解决方案	猿苑
缘圆	资源检索中的病毒防护	猿园
缘圆员	计算机病毒简介	猿园
缘圆圆	计算机病毒的特征和种类	猿员
缘圆猿	计算机病毒的防护	猿猿
缘圆源	重要的杀毒软件	猿缘
缘圆缘	反计算机病毒的相关网站	猿远
缘猿	综述与展望	猿园
缘猿员	基于 Internet 的化学化工资源检索	猿园
缘猿圆	基于 Internet 的化学研究环境	猿员
附录	邮件讨论组(电子论坛)清单	猿圆

1 Internet 化学化工信息资源的构成

员员 Internet 资源的构成

本节要点：

- Internet 的概念及其用途。
- Internet 资源有哪些？
- 如何确定 Internet 资源的位置？
- 浏览 Internet 资源所必需的软件。

员员员 员员员员资源的构成及其信息量分布

Internet 是近十多年内发展起来的全球性网络,分布于这一网络上的所有信息以及相关的各种服务统称为 Internet 资源。这些信息内容有各种文档、图片图像、声音视频、程序、数据库等,相应的服务有电子邮件、文件传递、网上教育、讨论组、联机文献检索、在线数据库、电子商务等等。关于 Internet 的发展历程,请参见 Robert Hobbes' Zakon 的 Internet 大事记 ([http ://www . isoc . org/zakon/Internet/History/HIT . html](http://www.isoc.org/zakon/Internet/History/HIT.html)),相应的中译文网址为 [http ://chin . icm . ac . cn/zakon/HIT-GB . html](http://chin.icm.ac.cn/zakon/HIT-GB.html)。

Internet 资源信息量大,内容覆盖面广,且呈爆炸式膨胀,仅化学化工专业信息资源就非常庞大,但相对于 Internet 的所有资源来说则如九牛一毛,要在 Internet 信息海洋中搜索到化学化工某领域的专业信息,就如大海捞针一般。因此,在检索 Internet 化学化工信息资源之前,学习一些检索方法非常必要,一旦掌握了检索方法和资源的规律特性,Internet 就会成为自己电脑中的虚拟化学化工图书馆。

Internet 并不单单是大家熟悉的万维网(WWW),它是由多种不同类型的信息资源共同构成的(见图 员员一个庞大的网络。只是由于近年来 WWW 资源发展极为迅速,远远超过了其他几种资源,使得 WWW 几乎成了 Internet 的代名词。而像 Gopher、Telnet 等曾经辉煌一时的东西已沦落为 Internet 的遗物了,从 Internet 各类信息资源流量的比例(见表 员员可证明这一点。

表 员员 员员员员各类资源流量比例

资源类型	Telnet	FTP	Email	WWW	其 他
比 例	员远%	愿源%	员园苑%	苑猷猷%	员%

为了响应 Internet 的发展趋势,本书主要论述 WWW 资源的检索方法。Internet 资源绝大多数为英文,一些经典权威的化学化工信息资源都以英文形式出现在西方国家的 Internet 站点上,这几年中文资源的增长也极为迅速,所以本书在主要介绍英文文化学化工信息资源的

1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 26

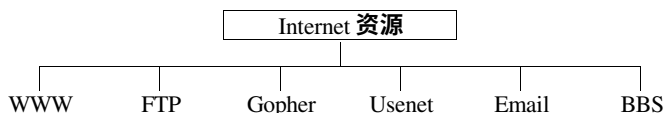


图 15-10 阿坝州主要资源的构成

Internet 资源极具动态性,随时都在发展和变化着,而且计算机专家们还在为它的繁荣而工作着,目前可以列举的 Internet 资源有如下一些主要方面,并且还在不断增加,每一个方面都蕴涵着大量的信息。因此,我们在学习 Internet 资源检索的一般方法时,还要不断地跟踪 Internet 资源的发展,以便更有效地利用它。

目前,Internet 资源大致归纳如下:

- **万维网** (WWW—World Wide Web)
- **电子邮件** (Email)
- **文件传输** (FTP——File Transport Protocol)
- **远程登录** (Telnet)
- **信息鼠查询** (Gopher)
- **电子公告板系统** (BBS——Bulletin Board System)
- **新闻讨论组** (Usenet/News Groups)
- **邮件讨论组** (Listserv/Mailing List)
- **查询 FTP 节点上的文件** (Archie Service)
- **查询 Gopher 节点上的主题信息** (Veronica 和 Jughead)
- **广域网信息系统** (WAIS——Wide Area Info Service, 可用关键词检索的分布式数据库系统)
- **视频会议系统** (CU-SeeMe)
- **音像点播和直播系统** (SteamWorks, VOD——Video On Demand)
- **适时广播技术** (AOD——Audio On Demand, Internet Wave)
- **IPX 仿真, 可以通过网络运行 IPX 协议的软件** (Kali)
- **用户查询** (Finger Service)
- **IP 地址查询** (Nslookup, Ping)
- **网际名址查询与检测** (Whios, Netfind, X. 纛纛等)
- **交互接力式对话或通信** (Chat, Talk 或 IRC——Internet RelayChat, Webtalk, Worlds Chat)
- **网络电话软件** (Iphone)、**传真、寻呼、PDA** (Personal Data Aid)
- **网际长途电话** (CoolTalk)
- **SGML** (Standard Generalized Markup Language, **标准通用标记语言, 是一种非常复杂的标记语言, 它的子集包括 HTML——Hypertext Markup Language 和 XML——Extensible Markup Language 等**)
- **JAVA 超文本开发语言**
- **CGI** (Common Gateway Interface, **这里的 CGI 包含网络上的 CGI、PHP、ASP 等运行于服务器上的程序**) **公共网关接口**

- SET(Secure Electronic Transaction)安全电子交易
- 加密软件(Key,PGP)
- Fire Wall 防火墙功能
- VRML 虚拟现实(Virtual Reality Markup Language)
- 多人网络游戏(MUD——Multiple User Dungeon、Multiple User Dialogue)

员圆 网络资源定位器——URL

Where? Internet 资源在哪里?这是我们检索这些资源之前首先要知道的。首先我们可以考虑一下这样的问题:我们在自己的电脑里输入一篇文章后,这篇文章你通常存放在哪里?你会回答:通常以文件的形式存放在电脑硬盘的某个文件夹中。好了,这篇文章就是一个资源,如果你的电脑接入了 Internet,同时你的电脑对 Internet 上的计算机提供服务,或者你把这篇文章放到了某台提供 Internet 服务的计算机上,那么这篇文章就成了 Internet 资源之一。从这里我们知道,Internet 资源其实就存放在所有与 Internet 连接的各个计算机主机(服务器)中,由于 Internet 上的主机(服务器)形式多种多样,数量以千万计,因此造成了 Internet 资源的宏大与复杂。

我们要查找某个信息资源,关键是要找到这条信息所在的计算机(主机)地址及这条信息在该计算机内的路径(通常是指 Internet 上的虚拟路径)和文件名。为了区别 Internet 上的每台计算机,连接到 Internet 上的计算机(包括提供服务的服务器和用于上网的客户机)都有一个惟一的地址作为标识,称为 IP 地址(通常使用四节数字表示,参见 员圆节)。而上面提到的每一个文件也都有一个确定其地址的东西,那就是 URL(Uniform Resource Locator 的缩写),专用于 Internet 资源的定位,它表示 Internet 上的某一个具体的对象,如一个文件或一个新闻组等,也就是通常所说的“网址”。其中文含义是“全球资源定位器”,就是它把我们带到资源所在的地方,并通过适当方式取得该资源。例如 `ftp://ftp.zdnet.nis.newscorp.com/pub/private/sw员/education/high_school_college/chemlb.zip` 的含义就是:请连接到 `ftp.zdnet.nis.newscorp.com` 这部主机,并利用 FTP 方式取得 `/pub/private/sw员/education/high_school_college/` 目录下的 `chemlb.zip` 文件,该文件是一个模拟化学实验的软件。

一个标准的 URL,其格式通常为: `method://hostname[:port]/pathname/filename`。其中 `method` 是指文件传输协议(资源获得方式),如 `http`、`ftp`、`gopher`、`file` 等;`hostname` 指服务器在 Internet 上的名称(主机域名),也就是我们通常所说的网站、站点或节点(注意 `method` 需要和“`://`”一起才能表示某种协议);`port` 指服务器提供服务的端口号,通常不输入,亦即使用默认的端口(`http` 使用 愿端口、`ftp` 使用 圆端口等);`pathname` 是指文件在服务器中的保存路径;`filename` 是指需要的文件名。

表 员圆中列举了目前网络中常用的 URL 类型、用途及示例。另外,用户也可以从使用的 URL 得知该 Internet 资源的实际类型,即从 `method://` 来判断资源是属于 WWW 类型或是其他类型。

表 员圆 常用 URL 类型、用途及其在 浏览器中的输入格式

URL 类型	用 途	示 例
Http	Hypertext Transfer Protocol,超文本文件传输协议,用于传输 HTML 及其相关文件类型	http ://www . chemworld . com . cn/
Ftp	File Transfer Protocol,文件传输协议,用于在服务器与其他机器间下载或上传相关文件	ftp ://ftp . tsinghua . edu . cn/
Telnet	远程登录协议,用于远程计算机控制	telnet ://locis . loc . gov/
Gopher	Gopher 服务传输协议	gopher ://gopher . seu . edu . cn/
News	新闻组或讨论组传输协议	news :sci . chem
Email	电子邮件传输协议	mailto :yuanzhongzhi @sina . net

员猿 检索 化学化工信息资源所需的软件

要获取 Internet 上不同类型的信息资源,通常应使用与各种资源对应的专用软件获取,但由于 WWW 资源发展非常迅速,且最利于 Internet 的大众化,获取 WWW 资源的软件大有把所有不同类型资源一网打尽的趋势。因此,常用的 WWW 浏览器软件(列于表 员猿中)除可以浏览 WWW 资源外,同样也可以获取表 员圆中所列举的其他非 WWW 资源(当然 WWW 浏览器本身应当包含相应的组件)。

表 员猿 常用的 浏览器

WWW 浏览器名称	说 明
Mosaic	NCSA 开发的世界上第一个成功的浏览器软件,微软的 IE 也是在此基础上进行开发的
Navigator 系列	Netscape 公司开发的世界上第一个成功的商业浏览器软件,被 AOL 并购后该浏览器软件由 AOL 负责继续开发
Internet Explorer 系列	简称 IE,Microsoft 公司开发的浏览器,从 IE源起已经成为浏览器的事实标准,目前最新版本是 IE缘缘功能和速度并重
Opera 系列	最小巧实用的浏览器,安装只需要一张软盘,对系统的资源消耗较少,支持大部分浏览器的重要功能
Lynx	DOS 下的字符界面浏览器,使用最低配置
Offline Explorer	功能强大和实用的离线浏览器,适用于下载感兴趣的整个网站内容,然后离线阅读
Teleport Pro	另一个著名离线浏览器软件,有不少特色和使用方便

为了保证读者能够最大限度地获取这些资源,本书所有的内容都是在普通个人计算机上使用最常用的操作系统和 WWW 浏览器获得的。根据国家网络中心和国内多家大型网络公司的网上调查,国内网络用户使用的操作系统中 Windows 怨缘或 Windows 怨愿网络浏览器中 IE源或 IE缘都占有 怨缘% 以上的比例,目前我们的建议配置是 Microsoft Windows 怨愿垣 Internet Explorer 缘园因此读者只需要掌握这两种软件的使用方法即可。当然,使用其他的网络浏览器也可以,如 Netscape Navigator。

图 员圆就是使用 IE 缘浏览 <http://www.yahoo.com/> 的界面,它支持表 员圆中各类 Internet 信息资源的获取,用户只需在地址输入框中输入相应的 URL(如果 URL 前没有写明相应的协议方法,则默认为 <http://>),然后按回车键(Enter)即可。如果该网络资源是存在的,浏览器软件就会自动将该信息资源‘拿’到你的电脑中并显示出来,你也可以选择保存到硬盘中。如果输入该资源时出现部分问题,浏览器软件一般会尽量尝试帮助你找到接近的资源,如果实在找不到则出现错误信息,最常见的错误信息是“服务器上没有该文档(Document Not Found)”、“找不到服务器”以及“内部服务器错误”,出现这些情况都表示你需要重新确认一下地址(URL)并做一些修改,有时也可能是因为对方服务器出现故障,需要稍后再重新链接。



图 员圆 浏览器界面及 URL 地址输入框

【思考题】

员 你所知道的 Internet 的服务功能有哪些?你使用过其中的哪些功能?使用过的浏览器软件有哪些?

圆 用于定位 Internet 资源的 URL 的格式有什么规律,如何记忆 URL?

猿 什么是 WWW 浏览器?列出你所知道的浏览器名称及其开发商名称。

源 通过帮助文件学习你所安装的 WWW 浏览器软件的使用方法。

缘 在 WWW 浏览器中如何输入一个 Internet 资源的地址(URL)并进行链接?例如 URL 可以为 ① <http://www.chemworld.com.cn>,或 ② <http://chemsoft.yeah.net>,或 ③ ftp://ftp.zdnet.nis.newscorp.com/pub/private/sw员b/education/high_school_college/chemlb.zip,或者其他 URL。

员圆 WWW 资源

本节要点：

- Internet 资源服务的 WWW 模式。
- WWW 域名与化学资源的关联。
- 常用 WWW 浏览器。

员圆员 什么是 宰宰宰？

WWW 是 World Wide Web 的简称,又称万维网、全球信息网或 Web 网,原本是位于瑞士日内瓦的欧洲粒子物理研究中心(CERN,European Laboratory for Particle Physics)的科学家们为了更好地共享信息而研制的网络工具,员愿愿年在 Internet 上公开后,整个世界立即被这个工具吸引,经过短短几年,特别是 员愿愿年伊利诺斯大学 Urbana-Champaign 分校的国家超级计算应用中心 NCSA(National Center for Supercomputing Applications)开发了 Mosaic 浏览器软件后,WWW 就成为 Internet 上的主流服务,大部分网站都立足于 WWW 服务,加上无数商业公司的参与,使得 WWW 空前繁荣,并具有无限发展潜力。WWW 目前由 W 猿(http://www.w猿.org)组织进行管理,并负责制定新的标准和内容。

WWW 的巨大成功主要来源于超文本(Hypertext)的思想,即页面上的文字、图片或其他任何对象都可以成为‘超链接’对象,单击后可以链接到 Internet 上的任何其他页面,信息之间是基于‘网状’形式组织的,可以相互参考,浏览无需按照固定的流程,真正实现信息之间的‘漫游’,因此人们形象地把在 WWW 中浏览和查找数据称为‘冲浪’(Surfing),不同主机或不同页面的信息犹如大海中的浪涛一样一阵一阵地扑面而来。同时,随着 WWW 技术的发展,WWW 上的信息不但可以包含文本、图形等,还可以包含大量的多媒体信息,最终 WWW 将发展成为一个跨越所有媒体和计算机系统的超媒体平台。

WWW 的传输方式是基于分布式客户机(Client)/服务器(Server)系统,按照 HTTP(Hypertext Transfer Protocol,超文本传输协议,WWW 服务程序所用的协议)进行文件传输和应答,而 HTTP 的基础是 Internet 最基本的协议——TCP/IP 协议(用于网络的一组通讯协议,包括网际协议 IP——Internet Protocol 和传输控制协议 TCP——Transmission Control Protocol),从而使得 WWW 成为 Internet 的一种标准模型。在 C/S 体系中,用户运行浏览器软件(客户),浏览器(如 Internet Explorer)通过 HTTP 协议获取 URL 所对应的服务器上的指定路径中用 HTML(Hypertext Markup Language,超文本链接标示语言)编写的超文本文件(即“网页”)或符合 HTTP 协议的其他类型文件,而 WWW 主机(服务器)则负责存储各种信息资源并运行相应的服务软件为用户的浏览器提供应答,传输要求的网页给客户机。

员圆圆 宰宰宰 域名与化学信息的关联

员 域名(Domain Name)

WWW 的信息是通过 HTTP 协议传递的,因此其 URL 特征为‘HTTP://域名或 IP 地址/路径/文件名’。这里‘域名’代表接入 Internet 中的一台服务器名称,例如,华南师范大学 WWW 服务器的域名为 www.scnu.edu.cn,实际上 Internet 上的每一个域名都可以使用惟一

对应的 IP 地址进行表示, www.scnu.edu.cn 的 IP 地址为 211.81.199.100 也就是说除非华南师范大学的 IP 地址或域名发生了改变, 否则两者在 Internet 上是等价的, 使用域名与 IP 地址的区别是域名容易记忆和理解而 IP 地址则速度较快。

服务器域名中的字母组合由多个层次表达, 层与层之间加“点”(Dot) 隔开, 从左到右其表达范围由小到大。如图 1-1 所示, 最右边的层为最高层, 通常代表国家和地区(依 ISO 标

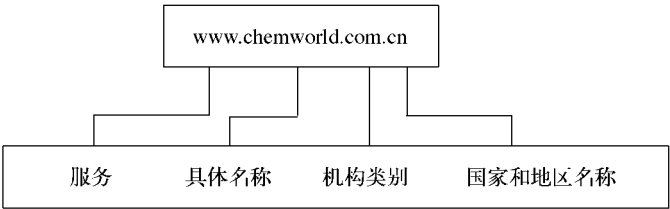


图 1-1 域名的含义

准用 3 个字母缩写表示, 如中国为 cn, 日本为 jp, 德国为 de, 等等), 但在美国的主机域名则没有这一部分, 即所谓的国际域名(注: 非美国机构也可以注册这种国际域名)。次高层限定主机所在机构的类别, 常用类别分类见表 1-1 源目目前还有一类直接表示地区的域名, 这类域名一般要与国家和地区名称组合才有意义, 如 .bj.cn 表示中国北京的某单位等)。再往左第三层才是主机所属机构或单位的名称, 它具有真正的含义, 人们争相注册的就是这一个名称, 它在 Internet 上是惟一的, 所以资源比较有限, 特别是简短和具有特殊意义的单词更是如此, 由于 Internet 管理组织规定在现实社会中已注册的商标在 Internet 上并不承认, 因此谁先注册该名称则使用权在规定的时间内就归他拥有。

表 1-1 域名中的机构类别及其含义

类 别	英文全拼	中文含义	类 别	英文全拼	中文含义
.com	Company	工、商业机构	.gov	Government	政府机构
.net	Network	网络服务机构	.ac	Academic	研究机构
.org	Organization	非营利机构	.mil	Military	军事部门
.edu	Education	大学和教育机构			

域名的机构名称部分一般由该机构名称的缩写或全称组成, 国内的域名则部分使用汉语拼音, 因此域名与 IP 地址相比具有更多的意义。例如美国化学会(American Chemical Society)的域名为: <http://www.acs.org/>, 国内如“羊城晚报”的域名为: <http://www.ycwb.com.cn/>。这一规律为我们记忆域名提供了方便。

1.2 域名与化学信息的关联

(1) 五花八门的化学化工域名。Internet 上的域名起得五花八门, 但域名通常是为了表达机构或组织的职能, 因此化学化工站点一般与化学(chemistry、chemical) 或它们的简写 chem(chemi) 有一定关系, 如果这几个单词(词根) 与其他相关的单词组合则可以表达新的含义, 这样就可以表达大部分化学化工机构的具体工作了, 注意这些信息可以作为机构名, 也可以作为主机名出现在域名中, 它们可以出现在图 1-1 所示的各部分中, 也可以分布于不同

的国家或地区。如化学与信息组合的域名 cheminfo, 可以有如下形式的组合: <http://www.cheminfo.com/>、<http://www.cheminfo.net/>、<http://www.cheminfo.edu.cn/>等。它们中 cheminfo 都是以机构名称形式出现的,而在 <http://cheminfo.nankai.edu.cn/>中 cheminfo 则以主机名(有时也称为次级域名)出现。

(圖 假设域名——查找 Internet 资源的另一种方法。由于域名与机构的职能和研究方向有关联性,因此如果你只有一个大概的兴趣方向,完全可以大胆假设域名,如果该域名已被注册并开通了相应的服务,一般都可找到与此相关的资源。现在我们来试试看,先试顶级域名,它最易掌握,而且一般主机在美国本土,历史较早,内容也较丰富。试试 chemistry 这一域名是否已存在,在浏览器地址栏中输入 <http://www.chemistry.com/>、<http://www.chemistry.net/>、<http://www.chemistry.edu/>或 <http://www.chemistry.org/>等,看是否有信息显示。若有,这些资料一般都会与 chemistry 有关(当然也可能出现错误提示,表示网络链接出错)。若没有信息显示,则有两种可能:一是这一域名未被注册,但一般较常用的单词或复合词都已被注册;另一种可能是虽已注册但未开通服务。试完国际域名后再试国家和地区域名。

下面我们再试试组合域名,如 chemistry 与世界(world)组合成为 chemistryworld、chemiworld、chemworld 等,chemistry 与软件(software)组成 chemsoft、chemistry-software 等,chemistry 与网络(web 或 net)组成 chemweb、chemnet 等,chemistry 与教育(education 或 instruction)组成 chemedu、chemcai 等,chemistry 与资源(information 或 resource)组成 cheminfo、chemsource 等。

总之,域名既然是为了易懂易记,一定会取一些清楚明了的短语,把 chem 与其他领域的单词(简写或全称)进行组合,一般都可得到相关信息。当然,要想使用这种方法来检索 Internet 资源无疑是碰运气,但若把它作为一种简单的单词组合游戏,有时会得到意外的惊喜。

(猿 中文域名。Internet 不是中国人发明的,域名当然不会考虑中文,但国内已有机构为了方便中国人使用 Internet,开发了中文域名系统‘猿’,通过使用类似虚拟域名或域名重新定向的方法,把用户输入的中文域名在专用服务器数据库中定位到真正的域名。“猿”的意思就是不管三七二十一,不管你输入中文或英文域名,都可在 Internet 上获取信息。关于中文域名的使用方法请参考 <http://www.猿.com/>。值得一提的是,使用中文域名实际比使用英文域名更有优势,如果一开始就使用中文域名的话,就不会出现今天的‘抢注域名’事件和域名资源缺乏的问题,因为在英文系统中好的域名(一般是一些非常常见和有意义的单词或单词组合)实在太少,而相应的中文则可以千变万化,中文在表达意思方面比英文更复杂和细腻。

WWW 资源已成为 Internet 上资源的主流,关于这一部分资源的检索将在以下各节逐步介绍。

【思考题】

员 写出你所在的单位的 WWW 服务器的 URL,或写出为你提供上网服务的网络服务提供商(ISP)的 URL,并解释 URL 的含义,哪一部分是 WWW 域名?

圆 WWW 域名的格式有什么规律?如何记忆这些域名(网址)?

