

第一章 化学信息学与 Internet

第一节 化学信息的构成及化学信息学

1. 什么是化学信息

什么是信息 信息是“含义”是一种“代码”，它被用来精确地描述物体或事件（以消除不确定性）。信息是独立于物质和能量以外的自然的第二属性。例如，生物基因包含了物种遗传所需的生物信息，语言是人们用来表达思想、情感和传递知识的人类信息。计算机和网络通讯中的二进制编码是现代信息技术的基础。当今世界是信息社会，信息已经成为与材料、能源同等甚至更为重要的资源，成为促进和发展科学、教育、政治、经济的基础。

化学信息包括化学化工及其相关领域的科学研究、化学工业生产技术、投资与商务活动、教育与教学过程中所涉及到的化学结构的信息、化学与化工反应信息、图谱与物理化学性质等数值信息、化学文献信息、专利、化学化工商务信息、远程教学、远程会议与在线交流、网络上的其他信息等等。更为广义的看法认为化学信息应该包括化学过程中分子间的相互适应、相互识别，例如在超分子的化学形成过程和分子自组装过程中所表达的分子信息。化学信息学是化学、化工与信息科学、计算科学的交叉学科，它涉及化学化工信息的获取、管理、处理与控制、计算推演与模拟和图形表示的技术与方法。学习化学信息学的目的就是要培养学生的信息意识和提高获取信息的能力。

化学是一门古老的科学。在很长的发展历史中，积累了大量的实验事实、数据与文献。在中世纪，炼丹术的秘方是由师傅通过口授方式传递给他的徒弟的；15世纪以后，化学信息开始以教科书的方式传递；从19世纪开始，科学技术的飞速发展引起信息的大爆炸；近50年来，新的实验方法、实验技术所产生的信息量极其浩繁，化学文献更是以每年数倍的速度迅速增加。近代科学的显著特点之一，是学科之间的相互渗透。化学家为了进行科学研究，所需要的信息范围甚广，所涉及的学科范围很宽。随着以计算机与网络为标志的现代信息技术的迅速发展，化学信息的数量日趋增加，形式也呈现多样化。

今天，被称为信息高速公路的 Internet（因特网）上存在着数以百万计的计

计算机主机,Internet 上的信息就分布在这些世界各地的主机上,信息资源极其丰富 约占全部信息资源的 20% 以上,随着图书情报界、图书出版界进入网络化、电子化和图书馆虚拟化的时代,这个百分数正在迅速增加。Internet 是当今世界上最大的信息传播媒介,它把世界的各个大学、机构和研究室联系在一起,把学习者、教育者、研究人员联系在一起,共享各种信息资源。World Wide Web 极大地影响着化学信息学的形成和以后的发展。在线文献检索、在线数据库检索、在线交流、远程学习和远程计算都已经成为现实。因此,在信息时代,对化学文献、化学数据仅仅以传统的方式处理和传递就不会适应实际的需要了,必须学习运用计算机管理、处理和表达化学信息,通过互联网获取与传递化学信息。

获取化学信息的三个基本途径是:检索、阅读文献;实验测量、查阅数据或图谱、数据处理;根据已有化学知识和物理定理通过计算机推演、大规模计算预测与模拟或再现化学结构、化学性质和化学过程。这三个基本途径效率的提高大都得益于计算机技术和互联网通讯技术的高速发展。

2. 化学信息学领域范围

化学信息学曾经被狭义地定义为化学文献学,主要是指化学、化工及其相关领域文献的检索与利用,而计算机与信息技术的发展和日益普及大大地拓展了它的内涵。现在,化学信息学 (Chemical Informatics) 是以化学、化工、计算机与信息技术为基础的边缘学科,并涉及生物、材料、环境、能源、地球与空间资源、冶金、器件 (传感器、分子器件与纳米机器) 等相关领域,它包括以下六个方面:

- 化学、化工文献学:传统方式的和电子与网络时代的文献信息检索与个人资料管理;
- 化学知识体系的计算机表示、管理与网络传输:化学结构、化学反应的计算机表示,化学数据库技术,化学信息的网际通讯语言;
- 化学图形学:化学信息可视化和虚拟技术,化工制图;
- 化学信息的解析与处理:化学实验设计,实验数据处理,图谱的分辨与解析,生物分子的信息解析,多元分析与数据挖掘技术;
- 化学知识的计算机推演:结构与性质关系,分子及其聚集体系的计算与模拟,分子与材料设计,化学反应的分析与设计,化工过程计算与仿真,专家系统;
- 化学教育与教学的现代技术与远程信息资源。

3. 化学信息学的成就与发展

化学信息科学在近年来最值得注意的成就之一是化学结构的计算机编码和图形检索方法的发展。传统的文献信息产品经过发展与转化后无一不可实现电子化、网络化 例如 美国《化学文摘》(CA)的光盘版 CA on CD、《科学引文索引》

(SCI) 的网络版 Web of Science 和 Chemistry Server 等。

化学计量学中的多元统计分析、人工神经网络、遗传算法、小波分析等先进的数据挖掘技术在化学的许多领域都有成功的应用。

今天化学领域的知识并非都来自于传统实验室的经验测量，其中部分是根据已有知识和定理通过计算机推演、大规模计算预测与模拟得到的。美国化学会在期刊 Journal of Chemical Information and Computer Sciences (化学信息与计算机科学) 上增加了副标题 “Includes Chemical Computation and Molecular Modeling” (包含化学计算和分子建模)。化学信息学正在设法提供一个以用化学结构为框架的通用化学语言来组织化学领域的全部知识。

近年来，可视化和虚拟技术、分子建模和计算化学、分子的物理性质和光谱技术、化工仿真技术取得了巨大成功，影响着化学、化工信息产品的发展，例如 ChemOffice、Cerius²、WebLab Viewer、Gaussian、MOPAC、AspenPlus、DesignIt、WinSim 等等。有些软件在 Internet 上甚至可以免费得到。典型的化学信息产品有：网络与文献信息产品，分子建模图形与设计软件，化工仿真与工艺设计软件，化学化工专家系统，化学化工数据库（化学与物性数据库，化学反应数据库，晶体结构数据库，图谱数据库，基因、蛋白质等生物大分子数据库），各种化学应用软件，化学教学课件与远程课程等。

化学信息学将在以下领域取得进展：

化学 3D 结构的计算机表示、存储与检索；

合成反应的自动分析；

化学信息的网络传输；

大型化学化工数据库；

与数据库相结合的化学化工数据挖掘技术；

高效率的非局域优化技术；

化学信息的可视化技术与虚拟技术；

分子动力学从头算方法；

生物大分子建模、构象分析与第一原理计算；

基因、蛋白质分子的生物信息解析；

溶液结构；

分子器件、纳米组装的模拟与计算；

材料结构与物性；

相图计算与化工过程仿真技术；

远程教育系统。

我国的现代计算机化学与技术研究起步于 20 世纪 70 年代，南京大学忻新泉教授运用计算机开始在数值处理和非数值运算方面展开工作，并率先在南京

大学开设了计算机化学类课程，推动了我国计算机辅助化学教学的研究活动；中国科技大学张懋森教授开展计算分析和专家系统领域的研究工作，并成为 1984 年我国的第一本计算机化学期刊《计算机与应用化学》的主编；中国科学院上海冶金研究所陈念贻教授将化学模式识别和计算机模拟方法应用于材料研究；中国科学院北京化工冶金研究所许志宏教授提出建立化学化工数据库。国内还有许多学者在计算机化学和计算化学的重要领域中作出了贡献：中国科学院上海有机化学研究所郑崇直教授、湖南大学俞汝勤教授、中国科学院长春应用化学研究所许禄教授、中国科学院药物研究所嵇汝运教授、中国药科大学安登魁教授、清华大学胡鑫尧教授、北京化工大学麻德贤教授、南开大学林少凡教授等。

第二节 因特网 (Internet) 概述

Internet 是由异型计算机、异型系统、异型网络系统协议所组成的联系世界 150 多个国家和地区各种大小计算机网络的互连体，正像其英文名称所表现的 Internet = Interconnect + Networks。

20 多年前，计算机只能在高规格的实验室看到，Internet 还只是美国很少的几台大型计算机连成的网络，使用者只限于个别国家的小部分研究人员。在那个年代，“个人计算机”这个概念都显得很荒谬，更不用说把无数台计算机连成一个全球网络了。科学发展史表明，任何一种成功的新技术对社会的贡献都在于它可以扩大人类的潜能。

今天的信息高速公路 Internet 具有在瞬间跨越万水千山交流信息的能力，使得人与人之间的距离缩短了，世界变小了。一进入 Internet 即可领略到其信息海洋之浩瀚：在 Internet 上可以轻而易举地到达世界各国的站点，到达许多公司、机构、大学的计算机，浏览或索取他们公开的软件或信息；可以参加各种各样的新闻组或电子论坛；可以发布自己的消息；可以查寻几乎所有的信息；可通过主题、特定的关键词和词组或描述性语言来检索 Internet 站点上提供的数据库信息。

Internet 已成为一种全球性的文化现象，成为并列于电视、电话的第三大通信系统。Internet 中的各个子网络由不同的国家管理，但又不受国界的限制，它们在网上维系着各自不同的风格和习俗而又协调地交叉在一起。现代信息社会把信息作为最重要的资源和产业，它和材料、能源一起并列成为社会物质和精神活动的三大重要资源。据估算，有效地利用信息、合理地提取其中的有用成分能产生数倍的附加经济效益。目前只是处于 Internet 时代的黎明，它的发展将对社会进步、科技发展、经济增长等起越来越重要的作用。信息技术已成为世界各国在科技战线上争夺的战略制高点。

第二章 化学化工网站与网上数据库

化学信息的需求是多方面的，除了大量的化学专业信息以外，还会涉及物理、数学、天文、地质、环境、计算机技术、工业技术、市场、投资、金融、法律、教育。我们应根据不同需要，掌握一定的检索技巧，灵活运用各种检索策略，学习在 Internet 上快速、准确、全面地获取信息的方法。

Internet 上的 WWW(World Wide Web) 环球网是一个可以检索 Internet 上几乎所有信息资源的信息查询系统，它利用超级文本的格式将相关联的信息文件有机地组织在一起。WWW 中的超文本(Hypertext)像 Windows 中的帮助文件，其中每个关键词都被用高亮度的色彩显示，但它的功能又远远超过帮助文件。WWW 的超文本文档形式不仅限于文本，而且是图文并茂、色彩丰富并可带有声音、动画等影视信息的超媒体，具有超出常规描述的多种传播媒体所具有的功能。这些文档往往带有超级链接(超级链接就是用醒目色彩的文字或图片链接着与它们有关的背景材料和参考文章或图片等，简称超链接)，使得浏览者只需轻轻地点击它就可以非常方便地进入被链接处，可以从一篇文章转移到相关内容的另一篇文章、从一台主机转到另一台主机浏览下去。

Internet 的发展十分迅速，一方面信息非常广泛而且很多信息在定期或不定期更新，另一方面信息资料的混乱度也在增加，信息的平均质量严重下降，系统化整理成本越来越高。人们不得不花大量宝贵时间去阅读价值不大的资料。因此，信息质量的重要性远胜过数量，信息挖掘的深度胜过广度。Internet 信息检索工具尤其是第二代检索工具可以帮助用户系统地浏览所要检索的目标，筛选和提高信息检索的速度和质量。

Internet 上的化学资源非常丰富，除了有提供通用信息检索的网站以外，还出现了针对化学领域的综合性信息服务和各种化学化工专业的站点。它们有些提供文献和化学化工信息检索，有些提供各种化学数据库，有些是化学论坛，还有些是为化学品、化学仪器、化学软件采购提供帮助。

化学数据库包括化学文献数据库、化学结构数据库和化学物质的物性数据库。数据库不仅提供原始数据，也是进行信息再加工的数据源，在科学研究和工程技术领域有着非常重要的作用。功能强大、内容丰富的大型数据库应用范围广，而小型、实用的专业数据库价格便宜使用方便。随着 Internet 的普及与发展，许多数据库提供网上查询服务。化学文献数据库在其他章节中介绍，本章中的数据库主要是指结构数据库、反应数据库和物性数据库。

第一节 通用信息检索引擎

Internet 检索工具一般具备以下功能：首先有一个简单易用、功能强大的前端界面，使得查询方式容易设定；检索方式比较完备，可以按用户所需的方式进行查询；检索结果具有较好的准确性和可读性，且检索速度较快。网络检索工具大都采用 Server/Client 结构和交互式的 Windows 前端界面，检索途径多，可使用“书签簿”保留检索历史，在检索窗口中按一定规则输入检索式，发出检索命令后即可得到检索结果。检索结果通常包括题目、内容简介、URL 地址等，由此可进行更进一步的检索，获得更详细的内容。由于检索引擎能够帮助用户在比较少的时间里方便地找到所需要的信息资料，因而成为用户上网的基本辅助工具。

WWW 是 Internet 上最先进的网络信息检索系统，它把超文本技术、网络技术和多媒体技术融为一体，并把 Internet 上的信息按一定的规则组织起来，以提供进一步的查询。从 1994 年 Yahoo 问世以来，已经开发出许多 WWW 检索引擎，有记录可查的检索引擎数量就达到 2 500 个左右。其中既有大型综合型的通用信息检索引擎，也有在特定领域里发挥作用的专业检索引擎。检索引擎的功能也有不少发展，第一代检索引擎是根据关键词在主页中的位置和出现频率来计分和排序，而第二代检索引擎以用户访问该网站时所花费的时间来确定它的人气指数，或者根据该网站被其他网站链接数量来确定它的等级，并向用户返送这些信息。

1. Yahoo(<http://www.yahoo.com>)

Yahoo 主页见图 2-1。Yahoo 提供按主题词分类的索引，并且有功能完备的检索方式，非常有效，但也异常拥挤，每天有上千万人次访问该节点。它已经开发了很长时间，由很多编辑人员维护，所以质量非常高。Yahoo 是目前最常用的检索工具，当不知去哪里时，在 Yahoo 上检索是找到好站点的最好方法，初学者也可以从那里找到较专门的站点。

使用 Yahoo 检索非常简单，输入主题词后点击“Search”按钮，或者点击“Advanced Search”（高级检索）即可。它将返回三种信息：

- (1) 满足检索条件的 Yahoo 的目录（用户可以利用这些目录进行交叉引用）；
- (2) 满足检索条件的 WWW 网站；
- (3) 含有网页索引的 Yahoo 目录。

Yahoo 提供了全面的分类体系，它将 WWW 服务的节点按主题词建立分类索引。我们还可以按这个分类来检索，检索的方法是先选出分类的超链接，在

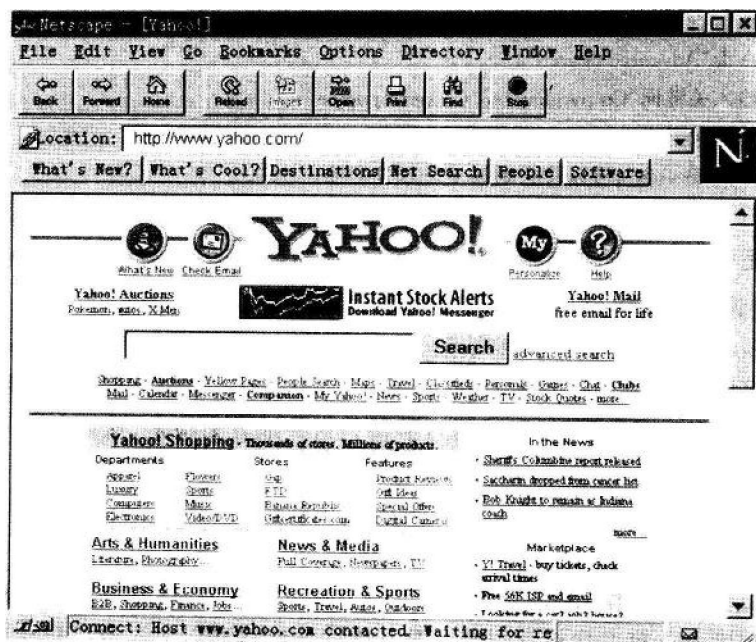


图 2-1 Netscape 浏览器窗口图 (显示内容为 Yahoo 主页)

Yahoo 查询结果页面中具有下划线的均为 Yahoo 分类的超链接,如“Business & Economy”(商务与经济)、“Regional”(地区)、“Education”(教育)、“Science”(科学)等等在选择分类的超链接后,从 Yahoo 就会得到一个查询界面,此时可以按其查询要求进行进一步查询

Yahoo 还提供一些简单的高级查询,不用通过分类索引直接检索关键词利用这种方法用户可以省去很多分类检索的步骤,但是注意使用的关键词的含义要相对窄一些。例如,对“chemistry”的检索,会使你面对数万条检索结果失去了检索的意义

对输入的关键词, Yahoo 提供 4 种检索方式和 2 种检索范围。它们是: “Intelligent default”(默认的智能方式)、“An exact phrase match”(精确词组匹配方式)、“Matches on all words”(AND, 单词之间是“与”关系)、“Matches on any word”(OR, 单词之间是“或”关系);“Yahoo Categories”(在 Yahoo 的分类网页范围内检索)、“Web Sites”(在 Yahoo 提供链接的 WWW 站点上检索) Yahoo 还提供在 Usenet 上检索其中“Intelligent default”是将输入的词作为一个子串,例如输入“meth”,那么 Yahoo 将检索 methyl(甲基)、methamphetamine(脱氧麻黄碱)和 method(方法)等。点击“Advanced Search”超链接,可以进行检索方式和检索范围的选择设置。

例 2-1 检索关于 Cu 与吡啶(pyridine)的反应的信息。

- (1) 点击 "Advanced Search";
- (2) 输入关键词 "Cu pyridine" 或 "Copper pyridine";
- (3) 将检索方式设为 "Matches on all words(AND)";
- (4) 点击 "Search", 可以检索到数千条有关信息。

在本例中, 如果选择关键词 "Copper pyridine reaction", 缩小检索范围, 检索结果的词条还可以少一些。对检索一些不能分开的词组如 "Organic Chemistry", Yahoo 规定该词组用双引号 "" 标记 即输入 "Organic Chemistry"。

Yahoo 在高级查询上有些缺陷, 如没提供 "NEAR" 等。但是 Yahoo 在其主页的末尾提供了到其他检索引擎如 Altavista、Lycos 等的超链接, 如果用户在 Yahoo 上查不到所需资料, 可以通过链接到这些地方去检索, 从而起到一定的补充作用。

2. 其他检索引擎

第二代检索引擎: 网站的质量反映在它内容的丰富性、准确性以及满足用户需求程度上。第二代检索引擎以被其他网站链接和以大多数用户访问的情况认定一个网站的质量, 它们的代表有

<http://www.google.com>;
[http://www.links2go.com\(Link To Go\)](http://www.links2go.com(Link To Go)) ;
<http://www.DirectHit.com>。

精选网站检索引擎: 它们提供高质量的网站, 对所挑选网站的质量以星的多少来表示。一个网站的星越多, 则其内容越好。

<http://www.About.com>(About 让各行各业的专家推荐他们所在行业的最佳网站);

<http://www.eBLAST.com> (Britannica 不列颠百科全书精选网址);

<http://www.RefDesk.com/index.html> (Reference Desk 参考文献检索引擎)。

大型检索引擎 AllTheWeb (<http://www.AllTheWeb.com>, <http://www.ussc.alltheweb.com>): AllTheWeb 是目前 Internet 上收录网页最多的检索引擎 (约有 2 亿个) 他们计划收录 10 亿个网页。同时他们认为 AllTheWeb 的检索速度最快, 敢于标明用户每次检索所用的时间。

特色检索引擎

<http://www.Meta-List.net> (Meta List 收录有 20 多万个邮件列表);
<http://www.FindWhat.com> (提供产品与服务信息);
<http://www.InfoJump.com> (收录近 4 000 种报刊的 500 万篇文章)。

检索引擎的检索引擎: Internet 信息检索工具的特点在于它可以让用户通过关键词来查询网上的信息, 由于不同的检索工具有不同的组织与分类, 因此检索结果的表达方式也各不相同。现在一般用户已经无法知道目前有多少检索引擎, 为了方便使用, 一些检索引擎目录应运而生。它们专门收集世界各国的各种检索引擎。另外, 一种新的称为“网页检索器”的服务正越来越引起人们的兴趣, 用户只要把自己感兴趣的网页地址在“网页检索器”中登记, 并提供自己的电子函件地址, “网页检索器”就会经常地为该用户检索相关的网页, 一旦发现网页内容有新变化, 就用电子函件通知该用户。用户登记的网页地址的类型可以包括“http”、“FTP”、“Gopher”等。网页检索器可称为一位忠实的侦察员, 它可以及时通知你感兴趣的站点的内容变化, 以便于得到最新的信息, 这对定期跟踪某些站点的科技信息资源尤其有用, 结合检索工具, 合理地加以利用, 可以收到很好的效果。

<http://www.searchenginecolossus.com> (Search Engine Colossus 以国家字母顺序排列);

<http://www.searchengineguide.com> (Search Engine Guide, 检索引擎指南);

<http://www.searchpower.com> (Search Power, 分类排列数千个检索引擎);

<http://www.easysearcher.com> (Easy Searcher 2, 专业检索引擎目录);

<http://www.netmind.com> (Net Mind 网页检索器)

Internet 上还有许多其他优秀的信息服务商, 如:

Infoseek(<http://www.go.com>) 包含有许多 Yahoo 中没有的化学专业信息;

Excite(<http://www.excite.com/query.html>) 也是一个很好的检索引擎;

AltaVista(<http://www.altavista.com>);

新浪 (Sina)(<http://www.sina.com.cn>);

中文 Yahoo(<http://cn.yahoo.com>) 中文检索引擎;

OpenText(<http://www.opentext.com>);

WebCrawler(<http://www.webcrawler.com>);

Lycos(<http://www.lycos.com>);

美国在线 (AOL)(<http://www.aol.com>);

B to B(<http://www.b2bscene.com>) 电子商务网站;

Nightn(<http://www.nightn.com>).

第二节 重要的化学化工综合网站

1. 中国化学化工 信息资源 (<http://chin.icm.ac.cn>, <http://www.chinweb.com.cn>)

中国科学院北京化学研究所计算机化学开放实验室的 ChIN(the Chemical Information Network) 是国内主要的化学化工信息资源。它提供的分类资源如下：

- Chemical Companies 化学品、仪器设备和化学软件供应商的网站地址；
- Chemical Databases 各类化学数据库；
- Chemical E-Journals 化学类电子期刊；
- Chemical Information Providers 化学信息产品商和机构；
- Chemical Resources and Services in China 国内的化学资源；
- Chemical Software 各种化学软件；
- Chemistry Related Books 化学类新书介绍，网上书店，书评；

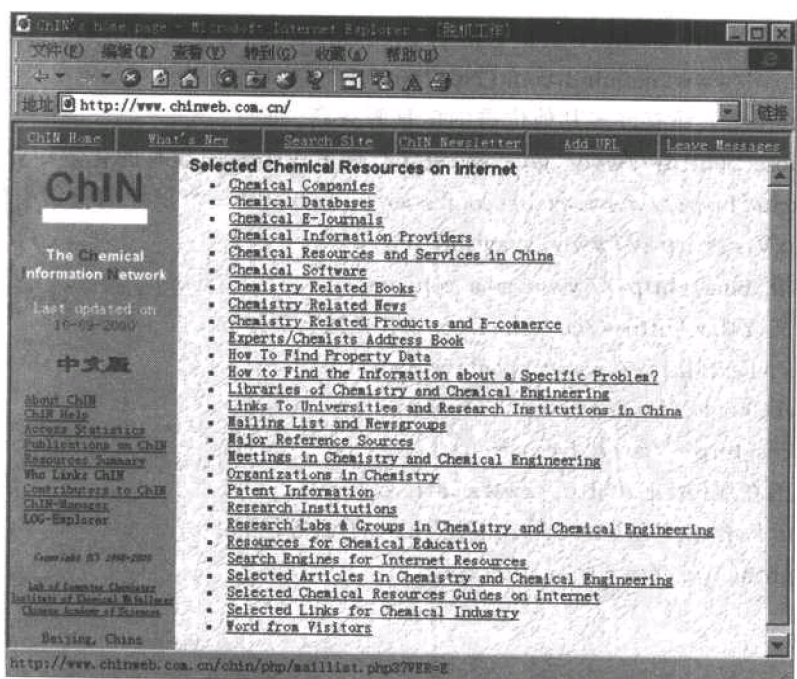


图 2-2 ChIN 的主页

- **Chemistry Related News** 化学类新闻；
- **Chemistry Related Products and E-commerce** 市场信息与供应商；
- **Experts/Cheemists Address Book** 工程类专家和一些研究人员信息；
- **How To Find Property Data** 物性数据查询；
- **How to Find the Information about a Specific Problem** ? 分类信息查询方法；
- **Libraries of Chemistry and Chemical Engineering** 化学化工图书馆；
- **Links To Universities and Research Institutions in China** 国内化学类研究单位和院系；
- **Mailing List and Newsgroups** 通讯讨论组和新闻组；
- **Major Reference Sources** 各种手册、大全、词典；
- **Meetings in Chemistry and Chemical Engineering** 化学化工会议；
- **Organizations in Chemistry** 化学组织与协会、政府机构、基金；
- **Patent Information** 专利信息；
- **Research Institutions** 国内和国外的研究机构；
- **Research Labs & Groups in Chemistry and Chemical Engineering** 化学化工实验室；
- **Resources for Chemical Education** 化学教育资源；
- **Search Engines for Internet Resources** 搜索引擎；
- **Selected Articles in Chemistry and Chemical Engineering** 化学化工热点文章；
- **Selected Chemical Resources Guides on Internet** 网上化学资源；
- **Selected Links for Chemical Industry** 化工企业。

2. ChemCenter(<http://www.chemcenter.org>)

ChemCenter 由美国化学会 ACS(American Chemical Society) 主办, 在该网站上美国化学会提供了大量的化学信息服务。从主页上(图 2-3)我们可以看到有大量的内容链接, 可以进行主题检索, 有 ACS 的会员服务。ChemCenter 的化学信息分类栏目如下:

- **Use the Library** 这个栏目主要提供文献信息服务, 是 ACS 的图书信息中心, 可以检索 ACS 的期刊出版物, 有当前的化学热点文章, 提供美国化学文摘服务社 CAS(Chemical Abstracts Service) 的付费服务；
- **Build Your Career** 提供继续教育与求职信息服务；
- **Find Meetings** 有美国化学会的全国性和区域性会议信息, 各分会和相关学科与领域的会议信息、讨论与论坛；

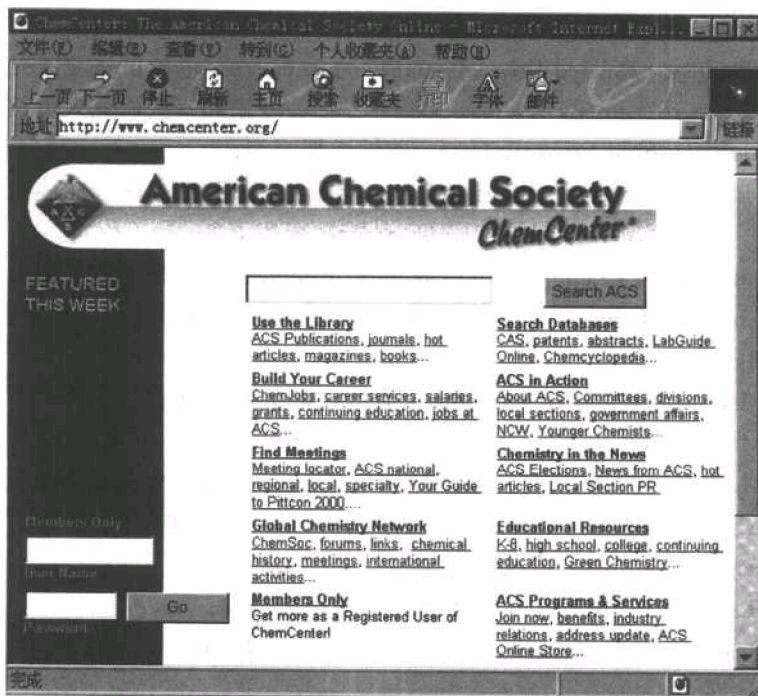


图 2-3 美国化学会站点 ChemCenter 的主页

● **Global Chemistry Network** 提供全球化学网页和网站的链接，主要有英国皇家化学会 (ChemSoc)、国际纯粹与应用化学联合会 (IUPAC)、各种论坛、化学会议、化学史、化学人物和国际活动，等等；

● **Members Only** 为美国化学会会员提供的服务，内含更多的信息资源，可以链接美国科学促进会 AAAS 的 Science 站点；

● **Search Databases** 提供美国化学文摘服务社 CAS 的 CA 和专利数据库检索，在线化学百科全书等；

● **ACS in Action** 介绍美国化学会的机构设置及各分会的活动；

● **Chemistry in the News** 各类化学新闻，其中有 ACS 的选举情况，ACS 的新闻，ACS 出版期刊中的热点文章 (Hot Articles)，ACS 的年报等；

● **Educational Resources** 介绍美国化学会的化学教育分会，有化学课程的材料，有职业教育信息，绿色化学资源。提供儿童的化学启蒙教育信息，高中与大专院校的化学活动和教材信息，研究生的专业目录和成人继续教育课程等；

● **ACS Programs & Services** 美国化学会计划与服务。

3. ChemWeb(<http://ChemWeb.com>)

ChemWeb 是个虚拟化学俱乐部, 它为化学家提供日常所需的网上信息资源, 包括 221 种化学类期刊, 32 个各类数据库, 免费的 AutoNum 命名数据库, 27.8 万种化合物, 90 多万种产品, 470 个供应商的各类数据。

网站上提供的服务有:

(1) 网上图书馆 (<http://chemweb.com/library>) 有 Marcel Dekker 公司提供的在线化学和物理类期刊。任何人可以免费查阅所提供期刊的目录与文章的摘要, 但是只有个人订阅户可以看到全文。

(2) ORGANIC CHEMISTRY FORUM 有机化学论坛 (<http://organic.chemweb.com>)。它提供大量有机化学信息资源, 如有机化学资源、个人的工作机会、会议详情、免费的有机化学期刊 Tetrahedron Letters 和它的预印刊 Tetrahedron Letters Fast、web 链接、数据库、讨论组、电子周刊 The Alchemist 的专题新闻和评论、ChemWeb.com 的在线杂志。

(3) BRETHERICK'S REACTIVE CHEMICAL HAZARDS DATABASE 在线危险品数据库 (<http://chemweb.com/databases/brethericks>) 付费服务。内含 5 000 多种化合物的数据, 可以用字符或结构在线查询。

(4) CHEMWEB 的免费目录服务 (<http://chemweb.com/library>) 为会员提供免费目录服务, 可以免费检索所有期刊和数据库目录列表。

免费数据库 <http://chemweb.com/databases>。

(5) ALERTING SERVICE (<http://chemweb.com/article-alerts>) ChemWeb 用 E-mail 提供更新通知服务。

(6) THE ALCHEMIST (<http://chemweb.com/alchemist>) ChemWeb 的 Web 电子周刊, 提供最新的化学新闻、化学评论及展望。

(7) THE LIBRARY (<http://chemweb.com/library>) 在线图书馆, 有覆盖化学的各个领域的近 200 种期刊。

(8) DATABASES (<http://chemweb.com/databases>) 有近 40 个各类数据库, 包括: 免费的可以由化合物化学结构图形自动生成 IUPAC 命名名称的命名数据库 AutoNom; 有机化学药物及同义词化学词典 NEGWER 合成化学词典。

(9) FORUMS 化学专题论坛, 各专题论坛组织了与该专业有关的化学资源, 其中有

分析化学论坛 (<http://analytical.chemweb.com>);

电化学论坛 (<http://electro.chemweb.com>);

催化化学论坛 (<http://catalysis.chemweb.com>);

有机化学论坛 (<http://organic.chemweb.com>)。

(10) CONFERENCE CENTRE 会议中心 (<http://chemweb.com/conference>) 免费向会员提供会议旅行服务 可以买机票、租车、预定旅馆房间。

(11) JOB EXCHANGE(<http://chemweb.com/jobs>) : 提供求职与招聘信息。

(12) SciQuest. com(http://chemweb.com/mall/mall_front.cfm?store=sciquest) 在线数据库 含有 650 000 种实验室设备和物品资源信息 , 可以检索供应商 , 比较价格 , 在网上完成购买。

4. 英国皇家化学会 RSC(<http://www.rsc.org>)及其 ChemSoc(<http://www.chemsoc.org>)

从英国皇家化学会(the Royal Society of Chemistry)的网站 RSC 主页(图 2-4)中可以看到,它主要有三个内容: PRODUCTS(信息产品)、ACTIVITIES(会议活动)、SERVICES(会员服务)。此外提供“SITE GUIDE”、“SITE SEARCH”、“ADDRESS BOOK”、“READ ALL ABOUT IT”四个链接。

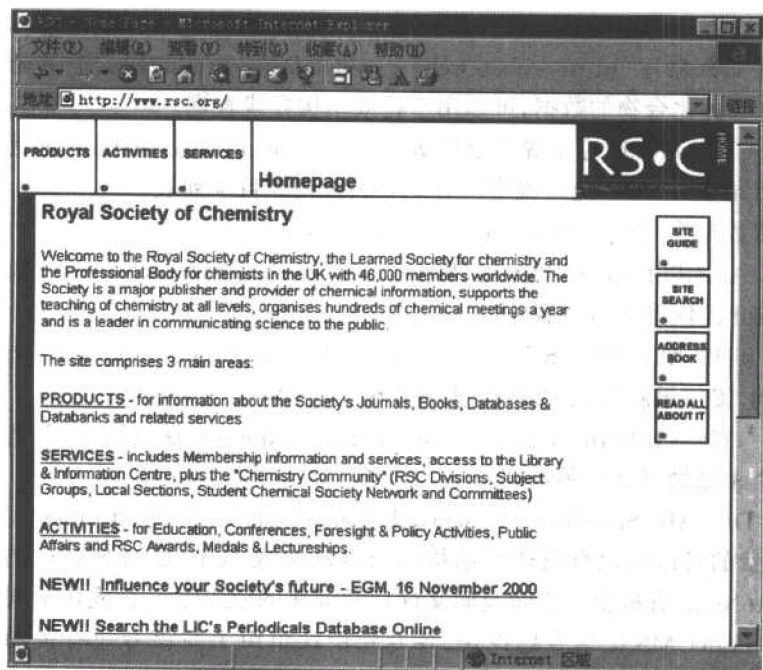


图 2-4 英国皇家化学会 RSC 的主页

(1)在 PRODUCTS 栏目中 Royal Society of Chemistry 介绍它们出版的化学类的书刊(journals and books)和数据库(databases & databanks)。

(2) **ACTIVITIES**提供教育、会议、各种活动的预告与通知 公布 RSC 的事务、奖励等。

(3) **SERVICES** 包括有会员信息与服务、图书信息、化学社区 Chemistry Community(“ RSC Divisions” ; “ Subject Groups” ; “ Local Sections ” ; “ Student Chemical Society Network and Committees” 的接入。

(4) **SITE GUIDE** 的链接。

(5) 在 **SITE SEARCH** 的链接中还提供当前 RSC 可供检索的数据库。数据库检索使用方法见“ Help with using Muscat”(<http://www.muscat.com>)。

(6) **READ ALL ABOUT IT** 热点新闻。

在 RSC 的化学社区服务页 ChemSoc(<http://www.chemsoc.org>) 中有更为个性化的链接。页面提供的链接按钮有：“Site guide”, “Site search”, “registration”, “web links”, “science park”, “what’s new”, “network & societies”, “careers and job center”, “conferences & events” ; “chembytes infozon” ; “109 visual elements periodic table”。

5. Chemdex(<http://www.chemdex.org>)

Chemdex 是英国 Sheffield 大学的化学站点。该站点最早建成于 1993 年 采用目录结构。现在它能提供 4 470 多个化学 Web 页的链接。页面安排分为几部分：按学科检索、按国家检索化学链接以及数据库等。此外还有化学学会组织、软件和其他一些有关化学的链接。

(1) 在化学学科分类中有：Analytical (分析化学), Biological (生物化学), Catalysis (催化), Computational (计算化学), Crystallography (结晶化学), Education (化学教育), Engineering (化学工程), Environmental (环境化学), General (普通化学), Geological (地球化学), Inorganic (无机化学), Laboratory (化学实验室), Materials (材料), Organic (有机化学), Organometallic (金属有机), Physical (物理化学), Safety (职业防护、化学品毒性与安全), Spectroscopy (波谱), Spectrometry (分析仪器), Theoretical (理论化学)。

(2) 大学和研究机构是按地区分类的，它们是：Africa, Asia, Europe, Germany, UK, North America, USA by state, Oceania, South America。

(3) 政府机关和实验室分别为 Agencies 和 Laboratories。

(4) 公司按地区 UK, USA, Rest of the World 链接。

(5) 在 Communication 中有：Journals, Publishers, Discussion groups, NewsGroups, Conferences。

(6) 数据库的链接中有 71 个数据库，以及用于教学的元素周期表站点 (<http://www.webelements.com>)。这 71 个数据库分别归入下列十三类：Bibli-

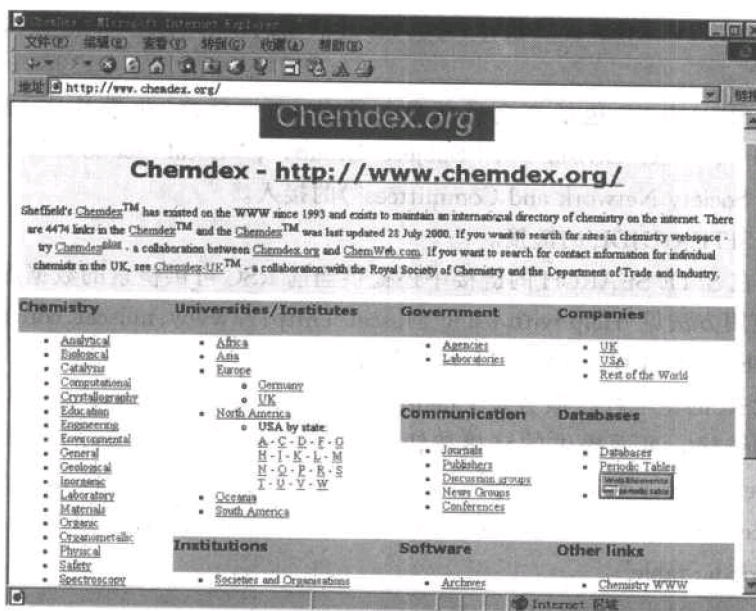


图 2-5 英国 Sheffield 大学的化学站点 Chemdex

ography search、Chemical Databases、Chemicals、Database on WWW、Education、Key Lists、Molecules library、Nuclear chemistry、Patents、Reactions、Reference Data、Reference Data via the WWW、Structure。

第三节 网上化学结构数据库、反应数据库和物性数据库

Web 数据库已经成为数据库网络化的一种基本形式，通过 Internet 可以非常方便地访问许多化学反应数据库和物性数据库。

1. CambridgeSoft (<http://www.cambridgesoft.com>, <http://www.camsoft.com>)

CambridgeSoft 提供 Internet 上的化合物和化学反应在线数据库，适合用于科学研究。此外，它还提供 ChemOffice WebServer，可以与其他的一些数据库相接，如药物类的 Derwent WDI、专利类的 Derwent WDA 和反应类的 ISI Current Chemical Reactions、InfoChem、ChemReact & ChemSynth 以及 Cambridgesoft 的 ChemRXN。

ChemACX (<http://www.ChemACX.com>) 是个很大的并且发展迅速的化合物商品供应信息数据库。它含有 Aldrich、Alfa - Aesar、Avocado、Fisher、Lancast-

er、Maybridge、TCI 等公司以及超过 200 家供应商提供的 175 000 种产量很大的化学商品和 9 家供应商提供的 250 000 种化合物的信息。此外，还有化合物的一些基本性质 如分子量、分子式和 CAS 登录号及商家提供的产品的其他数据，包括可供量、纯度、产品说明和定货信息。链接到供货商的 Web 站点能了解到他们的最新价格数据。

ChemMSDX 提供超过 7 000 种纯物质的物品安全数据的快速接入服务。

ChemINDEX(<http://www.ChemFinder.com>) 是一个很好的通用化合物性质数据库，全世界每天有数以千计的化学家使用它。其中包含了 NCI 公开的和许多新奇的化合物，有比 Merck Index 和 CRC Properties of Organic Compounds combined 更多的化合物和结构数据。它提供化合物的结构、CAS 登录号、分子式和相对分子质量 还可能有熔点、沸点、闪点、比重、蒸发率、蒸气密度、水中的溶解度、DOT 序号、RTECS 序号和 EPA 代码等性质和评价说明。

Derwent 的 World Drug Index(WDI)含有 58 000 种可用于药物或临床的生物活性化合物的数据。它可以在授权的局域网中用 ChemOffice WebServer 进行全结构检索。WDI 根据生物活性的类型将化合物分类，是有关上市和研发药物的权威信息源。数据信息包含结构、作用机理、同义词、商品名、制造商和说明。

Derwent 的 World Drug Alerts(WDA)提供新的生物活性化合物、新的药物合成方法和近期的期刊、会议、专利述评。它能在药物研发和生产全过程中，尤其对用于临床前候选药物的选择可提供决策帮助。ChemOffice WebServer 提供了通过授权局域网接入决策信息资源的理想方法。

Current Chemical Reactions(CCR 是美国 ISI 公司的化学反应数据库，有数据挖掘功能，能帮助进行设计化学合成反应。CCR 收集了超过三十万篇文章报告的分子的全合成方法。CCR 中的反应都由 ISI 的化学家验证过，并将结果返回给研究人员，然后将它们进一步分组和分类。

ChemReact and ChemSynth 由 InfoChem 提供。他们通过把反应分类，从一个含有 250 万种反应的数据库中挑选了 39 万种化学反应。ChemReact 帮助化学家设计新的合成方法。而 ChemSynth 中的反应是从 ChemReact 挑选出来的，其产率超过 50%，并且在文献中被报告过至少两次。

ChemRXN 是 Cambridgesoft 从 InfoChem、ChemSelect database 和 ISI ChemPrep database 中精选的反应数据库 内含 29 000 多个反应。ChemRXN 是 ChemOffice 的组件之一，可以在网上 <http://www.ChemFinder.com> 测试。

ChemOffice WebServer 具有子结构和反应检索能力，可以在 Web 上应用，是研究人员在局域网中接入基本数据源的理想工具。输入文件支持 SD 和 RD 格式，即用制表键(tab 键)定位的数据文件。ChemOffice WebServer 采用 ChemDraw Plugin 作为 Web 上化学结构式绘制与查询结果的结构显示平台。