

利用搜索引擎 检索现有技术

国家知识产权局专利审查协作中心 编写
魏保志 主编



知识产权出版社

全国百佳图书出版单位

利用搜索引擎检索现有技术

国家知识产权局专利审查协作中心 编写

魏保志 主编

知识产权出版社

内容提要

本书结合检索现有技术的特点,阐述了搜索引擎的应用技巧和实践:首先,介绍了综合搜索引擎及相关辅助工具的应用;其次,分类讲解了专利、学术、图书等专题搜索引擎;最后,介绍了在搜索引擎中建立检索域、进行追踪检索和特定技术内容检索等检索策略。

本书由国家知识产权局专利审查协作中心经验丰富的审查员编写而成,汇集了编写组多年的检索实践经验,同时结合了搜索引擎易于获得、技术先进的优点,对于提高广大读者的现有技术检索质量有一定的现实意义。

读者对象:专利审查员、专利代理人、信息检索员、科研工作者、工程技术人员以及相关领域的其他社会公众。

责任编辑:黄清明

责任校对:韩秀天

装帧设计:正典设计

责任出版:卢运霞

图书在版编目(CIP)数据

利用搜索引擎检索现有技术/国家知识产权局专利审查协作中心编写.

北京:知识产权出版社,2011.4 (2012.3重印)

ISBN 978-7-5130-0398-8

I. ①利… II. ①国… III. ①互联网络-情报检索 IV. ①G354.4

中国版本图书馆CIP数据核字(2011)第027908号

利用搜索引擎检索现有技术

Liyong Sousuo Yinqing Jiansuo Xianyou Jishu

国家知识产权局专利审查协作中心 编写

魏保志 主编

出版发行:知识产权出版社

社址:北京市海淀区马甸南村1号

网 址:<http://www.ipph.cn>

发行电话:010-82000860 转 8101/8102

责编电话:010-82000860 转 8117

印 刷:知识产权出版社电子制印中心

开 本:787mm×1092mm 1/16

版 次:2011年4月第1版

字 数:260千字

ISBN 978-7-5130-0398-8/G·384(3317)

邮 编:100088

邮 箱:bjb@cnipr.com

传 真:010-82005070/82000893

责编邮箱:hqm@cnipr.com

经 销:新华书店及相关销售网点

印 张:13

印 次:2012年3月第2次印刷

定 价:36.00元

出版权专有 侵权必究

如有印装质量问题,本社负责调换。

前言

科技创新和技术研发中的每一个环节都依赖于对现有技术的充分掌握,信息量剧增和知识更新换代的加速使得对现有技术的检索已经成为科研过程中的重要组成部分。

计算机检索已经引发了检索技术的第一次变革,提高了检索效率,并有力地推动了检索准确性的改进。随着科学技术和工程技术的进步,信息数据量的增速已经使单纯的计算机检索无法完全满足现有需求。搜索引擎拥有前沿的计算机检索技术,已经被应用在诸多方面,包括科研人员的文献调研过程等,并被证实为对于科技信息的检索具有极大的帮助。为充分开发利用搜索引擎检索现有技术,提高检索质量,国家知识产权局专利审查协作中心针对如何利用搜索引擎检索现有技术进行了专题研究,并将研究成果和检索经验集结成本书,以期帮助专利审查员、专利代理人、科研工作者、工程技术人员、信息检索员等掌握搜索引擎这一检索现有技术的有效工具。

本书分为6章,第一章“搜索引擎概述”介绍了搜索引擎的概念、种类和发展等,第二章“综合搜索引擎的利用”介绍了 Google、百度、Bing 等综合搜索引擎以及相关辅助工具,第三章“专利搜索引擎的利用”介绍了 Google Patents、百度专利等专利搜索引擎,第四章“学术搜索引擎的利用”介绍了 Google Scholar、Scirus、深度搜、期刊界等学术搜索引擎,第五章“图书搜索引擎的利用”介绍了 Google Books、百度图书搜索等图书搜索引擎,第六章“高级检索策略”介绍了“检索域”的使用、作者与引文追踪检索、特定技术内容检索等。第二至五章分别对应于对现有技术检索具有实际意义的几类搜索引擎,主要介绍了各类搜索引擎的基本使用方法、检索语法、检索技巧等,并辅以多个典型检索案例介绍了使用策略。第六章进一步介绍了高级检索技巧,其中“检索域”的建立通过对可检索资源的分类整合提高了检索的效率和准确性,而追踪检索和特定技术内容检索在现有技术检索中也具有实践意义。

本书由国家知识产权局专利审查协作中心审查业务管理与协调处组织编写,具体分工如下:

第一章由张辉、王扬平撰写;

第二章由吴获、马志勇撰写;

第三章由王扬平撰写；

第四章由李福永、马志勇撰写；

第五章由马志勇撰写；

第六章由张辉、马志勇撰写。

全书由周胜生、欧阳雪宇和刘慧进行统稿，由曲淑君进行全书审定。

在本书编撰过程中，部分案例采用了靳晶、马晓宇、何永春、张丹、刘中涛等人提供的内容，第六章的部分内容参考了专利审查协作中心化学发明审查处检索小组所编写的内部资料“Google 检索手册”，该小组成员包括孙瑞丰、袁海宾、李哲、王晓明、尹俊峰、唐志勇、李东博、易方、唐甜甜、张沫、张辉等，在此表示感谢！

由于作者水平有限，缺漏、不当在所难免，恳请读者批评指正。

本书编写组

2011 年 1 月

目 录

第一章 搜索引擎概述	(1)
一、搜索引擎的概念	(1)
二、搜索引擎的种类	(1)
三、搜索引擎的最新发展	(3)
(一) 搜索趋于智能化	(3)
(二) 搜索趋于个性化	(4)
(三) 搜索趋于多语言交叉	(4)
四、利用搜索引擎检索现有技术概述	(5)
第二章 综合搜索引擎的利用	(7)
一、Google 搜索引擎	(7)
(一) Google 基本检索与浏览	(8)
(二) Google 检索语法	(20)
(三) Google 检索技巧及注意事项	(29)
(四) Google 与 ISI 等非专利数据库的比较	(36)
(五) Google 检索辅助工具	(37)
二、百度搜索引擎	(41)
(一) 百度基本检索与浏览	(41)
(二) 百度检索语法	(46)
(三) 百度检索技巧及注意事项	(46)
三、Bing 搜索引擎	(48)
(一) Bing 基本检索与浏览	(49)
(二) Bing 检索语法	(54)
(三) Bing 检索技巧及注意事项	(55)
四、Google 与百度、Bing 的比较	(56)
五、检索案例	(57)
第三章 专利搜索引擎的利用	(62)
一、Google Patents 专利搜索引擎	(63)

(一) Google Patents 简介	(63)
(二) Google Patents 基本检索与浏览	(64)
(三) Google Patents 检索语法	(73)
(四) Google Patents 检索技巧及注意事项	(75)
二、百度专利搜索	(78)
(一) 百度专利简介	(78)
(二) 百度专利基本检索与浏览	(78)
三、专利搜索引擎与其他相关网站的比较	(80)
(一) Google Patents 与百度专利的比较	(81)
(二) Google Patents 与 USPTO 在线检索数据库的比较	(82)
(三) Google Patents 与 FreePatentsOnline 的比较	(84)
(四) Google Patents、百度专利与 Patentics 的比较	(86)
四、检索案例	(87)
第四章 学术搜索引擎的利用	(90)
一、Google Scholar	(90)
(一) Google Scholar 基本检索与浏览	(91)
(二) Google Scholar 检索语法	(101)
(三) Google Scholar 检索技巧及注意事项	(104)
二、Scirus	(112)
(一) Scirus 基本检索与浏览	(113)
(二) Scirus 检索语法	(117)
(三) Scirus 检索技巧及注意事项	(119)
三、深度搜	(119)
(一) 深度搜基本检索与浏览	(121)
(二) 深度搜检索注意事项	(128)
四、期刊界	(129)
(一) 期刊界基本检索与浏览	(129)
(二) 期刊界检索技巧及注意事项	(134)
五、学术搜索引擎的比较	(136)
六、检索案例	(138)
第五章 图书搜索引擎的利用	(141)
一、Google Books	(141)
(一) Google Books 基本检索与浏览	(142)
(二) Google Books 检索内容与检索语法	(146)

(三) Google Books 检索实例与下载技巧	(148)
二、百度图书搜索	(152)
三、Google Books 与其他图书系统的比较	(153)
第六章 高级检索策略	(157)
一、“检索域”的建立与使用	(157)
(一)“专利域”的建立与专利检索	(157)
(二)“非专利域”的建立与非专利检索	(169)
二、作者与引文追踪检索	(179)
(一)作者追踪	(179)
(二)引文追踪	(183)
三、Google 进阶检索策略	(185)
(一)Google 综合检索策略	(185)
(二)Google 对特定技术内容的检索	(189)
参考文献	(196)

第一章 搜索引擎概述

随着互联网的普及与发展，一方面，越来越多的信息由传统的纸件存储方式逐渐向电子化、数字化转变，互联网的信息搜索日益成为人们获取信息的主要手段；另一方面，随着计算机搜索技术的发展，信息搜索变得更方便、结果更可靠，人们也越来越依赖互联网的搜索。互联网搜索技术的集大成者便是搜索引擎，它是互联网搜索的重要工具，并随着网络的发展不断自我完善。

一、搜索引擎的概念

搜索引擎通常指的是互联网上专门提供查询服务的一类网站，用于搜索万维网（www）或文件服务器（FTP）的信息，将搜索得到的由网页、图像、声音等文件组成的结果以列表的形式显示，某些搜索引擎还具有数据挖掘、算法规则、基于用户输入的混合算法等功能^①。

上述定义集中反映了搜索引擎发展过程中的一段非常重要的历史。在万维网出现之前，基于分散的文件服务器的搜索工具，如 Archie、Gopher 等就已经出现，虽然它们只能通过文件名的形式进行搜索，但奠定了搜索引擎的基本工作原理，可称为早期的搜索引擎；另一方面，上述定义也反映了搜索引擎同时担负着网络信息资源的管理功能，需随时自动更新数据源、记录网页地址、建立索引结构等，以便迅速地响应用户的检索需求。

二、搜索引擎的种类

超文本传输协议（HyperText Transfer Protocol，HTTP）访问 Web 技术的日臻成熟，更多分散的文件服务器被连接起来，文件也不再局限于已有的格式，且各种数据通过互联网共享连接起来，如何在多样化的资源中准确地查询到有用的信

① http://en.wikipedia.org/wiki/Web_search_engine.

息,将更加依赖搜索引擎。此时以 Yahoo^① 为代表的搜索引擎建立了可搜索并可浏览的人工目录分类,它的出现适应了互联网网页数量以及用户访问量增长的趋势,且不再拘泥于特定的网络,而顺应互联网(或万维网)广泛应用后的现代搜索引擎,这种基于分类目录导航系统的检索方式,也被大多数人称为第一代搜索引擎^②。目录式搜索引擎除 Yahoo 外,尚有 Open Directory^③、LookSmart^④ 等。

搜索引擎发展的阶段划分各有不同,但通常均以 Yahoo 与 Google 的出现与成熟作为标志。例如,有人认为第二代搜索引擎是以 Google 为代表的关键词和特殊算法的搜索引擎^⑤,也有人认为 Google 是继 Archie、Yahoo 等之后的第三代搜索引擎^⑥。但无论如何划分,依靠机器抓取并建立在超链接分析基础上大规模网页搜索的 Google,将搜索结果的准确性与互联网信息的迅猛增长同步起来,配合其排序、多语言支持等功能,Google 被公认是目前世界上最流行的搜索引擎。与 Google 类似的搜索引擎还有 Lycos^⑦、OpenFind^⑧、Bing^⑨ 等。

上述以 Yahoo、Google 为代表的搜索引擎,按照工作方式的不同,又被相应地分为目录索引类搜索引擎和全文搜索引擎,另外还有一类为基于多个搜索引擎同时搜索后的结果,称为元搜索^⑩,例如 Infospace^⑪、dogpile^⑫ 等。

另外,根据搜索引擎的定位和其搜索范围,搜索引擎可以分为综合搜索引擎和各种专门搜索引擎。综合搜索引擎面向一般目的的检索,其检索结果包括互联网上的各种信息,一般不对检索范围作限定;但专门搜索引擎面向特殊目的的检索,常将检索范围限制在某一类信息中。以 Google 为例,其包括作为综合搜索引擎的网页搜索(Google Web Search),以及专利搜索引擎(Google Patents)、学术搜索引擎(Google Scholar)、图书搜索引擎(Google Books)等多个专门搜索引擎。

互联网上的信息浩瀚复杂,搜索引擎也纷繁多样,各种搜索引擎的搜索技术、网页索引页面数量各不相同,使用时应当注意选择。目前互联网上除上面已提及的搜索引擎外,按中英文资源的搜索重点不同,还有以下比较常用的综合搜索引

① <http://www.yahoo.com/>.

② 包瑞. 浅析第三代搜索引擎的发展 [J]. 晋图学刊, 2010 (4): 20-22.

③ <http://www.dmoz.org/>.

④ <http://www.looksmart.com/>

⑤ 包瑞. 浅析第三代搜索引擎的发展 [J]. 晋图学刊, 2010 (4): 20-22.

⑥ 李悦. 搜索引擎技术的产生与发展综述 [J]. 福建电脑, 2010 (5): 34-35.

⑦ <http://www.lycos.com/>.

⑧ <http://www.openfind.com/>.

⑨ <http://www.bing.com/>.

⑩ 包瑞. 浅析第三代搜索引擎的发展 [J]. 晋图学刊, 2010 (4): 20-22.

⑪ <http://www.infospace.com/>.

⑫ <http://www.dogpile.com/>.

擎（见表 1-1）。

表 1-1 其他较常用的中英文综合搜索引擎

序号	搜索引擎	网 址	类 型
1	搜狗	http://www.sogou.com/	中文搜索引擎
2	新浪爱问	http://iask.com/	中文搜索引擎
3	网易有道	http://www.youdao.com/	中文搜索引擎
4	百度	http://www.baidu.com	中文搜索引擎
5	alltheweb	http://www.alltheweb.com	英文搜索引擎
6	AOL	http://www.aol.com/	英文搜索引擎
7	altavista	http://www.altavista.com	英文搜索引擎
8	gigablast	http://www.gigablast.com	英文搜索引擎

三、搜索引擎的最新发展

互联网信息量的继续增长以及用户对搜索结果准确性要求的日益提高，搜索引擎的发展也随之朝智能化、个性化等多方向发展。在智能化方面，为使用户搜索意图与计算机理解之间趋于一致，出现了语义搜索引擎，例如 Kngine^❶、Swoogle^❷等；在个性化方面，信息将被进一步细分并建立独立的搜索引擎，例如针对图像的搜索引擎、学术的搜索引擎、专利的搜索引擎等。

（一）搜索趋于智能化

互联网上信息量逐渐增多，搜索的覆盖面也就越大，对返回结果的精度要求就越高，前面所述的以 Google 为代表的全文搜索引擎，虽然采用了网页排序算法等搜索技术大大地提高了检索的效率，但均以关键词的匹配为基础，同样无法回避一词多义、字符单纯匹配、缺少语义描述等缺陷，加上全文搜索的返回结果较多，搜索的准确性日益受到信息量“爆发式”增长所带来的挑战，因此人们日益重视搜索引擎的智能化程度。

搜索引擎的智能化将着眼于增强计算机的理解技术，加强计算机与用户之间的双向性沟通，使计算机更好地理解用户的搜索意图，并在搜索过程中以知识进行检索，而不再单纯以关键词的索引匹配方式进行检索。这种智能化搜索技术将

❶ <http://www.kngine.com/>.

❷ <http://swoogle.umbc.edu/>.

高度依赖自然语言的处理技术和人工智能技术^①，其中需要运用关键词的分词、同义词、反义词、短语识别、多语言翻译等技术来实现对用户检索意图的真正理解，另一方面，更重要的是需要进一步赋予搜索的内容（例如 web 网页）更多的语义信息。

互联网（或万维网）作为人们当前进行检索的重要手段之一，如何赋予互联网中的数据信息语义条件，形成具有良好定义的语义网^②，帮助人们高效快捷地获取互联网中的信息，成为语义智能化检索的研发先锋。语义网是相对下一代网络搜索引擎而言的智能网络，开发语义网主要是认为目前基于超文本标记语言（Hyper Text Markup Language, HTML）的互联网，缺少对检索文本中信息的描述，而语义网要实现的目的是，增加互联网的语义信息，提高机器的可读性，不仅能理解用户的检索意图，而且计算机之间能够进行有意图的沟通。

（二）搜索趋于个性化

个性化是指搜索引擎更注重搜索服务的个性化，针对用户的个性化需求进行定制并理解，实现搜索引擎的个性化模式，目的在于更有效地区分搜索对象，并准确地理解个性用户的搜索意图，更准确地返回搜索的结果。

目前搜索引擎的个性化研究主要集中于主题爬虫的个性化和查询改进等方面^③。基于查询改进的用户反馈技术是研究的重点，例如建立“学习型反馈技术”^④使搜索引擎更好地理解特定用户的搜索喜好和意图，将搜索结果按照学习理解后的相似度高低重新排序，优先返回高相似度的网页内容，以提高检索效率。

未来搜索引擎也将趋于专业化。专业化的搜索引擎将某一专题专门收录，不但减小了综合性搜索引擎的搜索压力，而且也为搜索引擎的个性化服务创造了有利条件。根据专业化搜索引擎的跟踪、分析、学习为用户定制更全面的专业搜索服务，使特定的用户与计算机之间的理解更加趋于一致，将极大地提高检索效率，这是今后发展的重要方向。

（三）搜索趋于多语言交叉

多语言交叉搜索在于跨语言信息搜索，指用户以一种语言（通常是自己的母

① 兰蔚巍，李海生．浅谈智能搜索引擎技术及其发展趋势 [J]．计算机与网络，2010：256.

② 崔亚军．语义网的起源与发展．机械管理开发 [J]，2009，24（5）：186-187.

③ 刘建生，周志辉．个性化搜索引擎综述．计算机与数字工程 [J]，2010（10）：80-81，94.

④ 江婕，李建民，曾劭炜．基于用户反馈的个性化搜索引擎的研究 [J]．计算机与现代化，2010（6）：116-118，121.

语)提问,检出另一种或几种语言描述的信息资源的信息检索技术和方法。跨语言信息检索中,用户用以表达自己的信息需求、构造检索提问式的语言称为源语言(Source Language),被检索的信息资源所使用的语言称为目标语言(Target Language)①。

近几年来,随着计算机技术的发展,机器翻译的技术也迅速发展。从传统的基于规则的机器翻译扩展到了基于实例或模版的机器翻译、统计机器翻译等,尤其是近年来语言学和人工智能技术的发展,以语义描述或以知识描述为特征的智能机器翻译系统正逐步成为研究的热点②。

搜索引擎对多种语言数据进行跨语言的交叉搜索,不仅能够返回用户搜索的所有语言的结果,而且随着机器翻译的进一步成熟,还能够进一步将返回的所有语言结果以母语的形式显示。虽然该技术目前还处于初步研究阶段,但其不仅符合搜索技术的发展方向,而且基于搜索引擎在未来语义智能化等领域的突破,跨语言交叉搜索必将成为现实。

综上,随着搜索引擎技术的进一步发展,以语义搜索为核心的技术推动了个性化、专业化、交叉语言搜索技术的发展,相信随着研究的进一步深入,在不久的将来,现存的语言障碍和搜索效率低下等缺陷必将逐渐被克服,不同层次的用户有望借助于智能化的搜索引擎实现专业化检索。

四、利用搜索引擎检索现有技术概述

现有技术是技术创新的基础,充分检索和合理利用现有技术是提升创新等级和降低创新成本的重要途径。专利制度是推动技术创新的有力保证,对于其中的专利审查环节,通过检索获得相关的现有技术还是判断发明或者实用新型是否具备新颖性或创造性的基础。

目前,现有技术的检索通常依赖于专门的检索数据库,这些数据库所使用的相关检索技术虽较成熟但略显陈旧,已成为制约检索质量进一步提高的因素。近年来搜索引擎发展迅速,传统搜索引擎不断扩展其功能,应用了大量先进的检索技术和理念,同时还出现了大量新的搜索引擎。为了进一步改进现有技术的检索,有必要探讨应用搜索引擎检索现有技术的可行性。以下综合评价搜索引擎的常规标准,从覆盖范围、查全率、查准率、响应时间、用户负担、结果输出方式等方

① 曹红兵. 跨语言信息检索在 Internet 搜索引擎中的应用 [J]. 江西图书馆学刊, 2006, 36 (3): 68-70.

② 周胜生, 王扬平. 专利文献计算机检索技术的最新发展 [J]. 图书情报工作, 2010 (18): 81-84.

面探讨搜索引擎在检索现有技术方面的应用前景。

就覆盖范围而言,由于检索所针对的专利技术大部分属于工程类信息,且可能涉及各个技术领域的内容,因而要求搜索引擎所覆盖的数据必须尽可能的全面。据经验来看,专利审查过程中所引用的现有技术主要来源于专利文献、科技论文,因而特别要求搜索引擎能够有针对性地覆盖这两类文献。

查全率和查准率关系到检索质量和难易。对于这方面而言,传统的目录式搜索引擎可能难以适应于科技信息的迅速增长,且人为加工的目录显然也无法全面覆盖各种信息。但目前主流的搜索引擎大多采用了自动分析页面和生成索引的方式,且已经通过各种技术对检索条件进行一定程度的智能分析、对检索结果进行智能排序,因而相对于传统的计算机检索数据库在查全和查准程度上并不逊色。需要注意的是,通用搜索引擎所覆盖的数据量庞大,需要注意检索过程中引入的非技术性信息噪声。

就响应时间和用户负担而言,一般搜索引擎的响应时间都在毫秒或秒的量级,并不会给用户带来不便,且互联网通用搜索引擎多为免费模式,这也给推广搜索引擎的应用带来了便利。

就结果输出方式而言,由于搜索引擎并非专门为了检索现有技术而开发,因而其检索结果输出方式仍存在一些不便,例如不便于快速浏览、图文并茂的浏览模式较简单、难以进行多次检索结果之间的逻辑运算,等等。然而,搜索引擎在结果输出方面的一些优点也是值得重视的,例如自动显示最相关的部分和高亮显示检索词,同时,一些通用搜索引擎也专门开发了涉及专利文献的检索工具,针对于文献特点优化了结果浏览,使其更适合于现有技术的检索。

本书将介绍如何利用搜索引擎检索现有技术,所涉及的搜索引擎包括 Google、百度、Bing 等搜索引擎以及其所提供的专利、学术、图书等专题搜索引擎,通过综合对比并结合案例分析介绍了各类搜索引擎在检索现有技术过程中的使用技巧。

第二章 综合搜索引擎的利用

综合搜索引擎是一种集成了多种搜索产品的综合性搜索平台，可从其所收录的丰富的网络资源中为用户检索到所需的信息。综合搜索引擎根据各种可检索的网络资源的用途进行分类，并针对每种类型的网络资源提供有针对性的检索，从而形成多个各有特色的搜索产品，例如网页搜索、图片搜索、地图搜索，它们的检索对象分别是网页、图片、电子地图。用户进入综合搜索引擎后，可以根据自身的检索需求，选择一种搜索产品进行搜索。

网页搜索是应用最为普遍且资源最为丰富的搜索产品，虽然它不是专用于检索技术资料的，但是仍然能为技术人员提供各类有用信息。例如，从高校、科研院所或研究人员的主页中能追踪当前的研究热点和某人的科研活动，从技术论坛、技术领域的门户网站上能深入了解某一个技术领域的发展现状、最新产品，甚至并未收录在学术文献数据库中的技术文档，从制定业内标准的网站上能了解某项技术标准的关键技术、演进过程，从专利搜索网站上能定位到相关专利文献。

目前应用较为广泛的综合搜索引擎包括 Google、百度、Bing（必应）、Yahoo、AOL、Global Spec，等等。其中具有较大影响力的 Google 能够搜索的外文网页和文献最多。百度是国内最大的中文搜索平台，无论是用户量还是所开发的产品数量都处于前列。Bing（必应）是微软公司于 2009 年推出的搜索引擎，由于上线时间较晚，目前还不是很成熟。

本章将分别介绍 Google、百度和 Bing（必应）这三大搜索引擎的主要功能，以及它们在检索现有技术方面的应用技巧，并简要分析三者之间的异同。对于每个搜索引擎，将主要介绍网页搜索功能，并辅助介绍与技术资料的检索密切相关的特色搜索产品。

一、Google 搜索引擎

Google 检索系统提供了许多搜索产品，最常用的为网页搜索，即 Google Web Search（英文版）。Google 网页搜索的特点包括：

- （1）收录了数以亿计被 Google 抓取的各类网页资源，包括门户网站、高校网

页、个人主页、博客、论坛等；

- (2) 能搜索多达 51 种语言的网页，并提供不同语言的网页之间的翻译；
- (3) 使用独有的 PageRank™ 算法对搜索结果进行排序；
- (4) 提供简单搜索和高级搜索两种不同的搜索模式；
- (5) 提供多种检索命令和字段来灵活配置检索表达式；
- (6) 提供百宝箱工具的多种选项用于过滤搜索结果；
- (7) 提供“手气不错”功能直接导向最接近的网页；
- (8) 对于同一学术文献可提供比 Google Scholar 更多的下载链接地址。

(一) Google 基本检索与浏览

本节将详细介绍 Google 网页搜索提供的检索入口、更多检索服务、结果浏览页面以及百宝箱工具。此外，本节还对 Google 可检索的专利和非专利资源进行了总结。

1. 检索入口

通过 Google 主页即可使用 Google 检索系统，主页地址为 <http://www.google.com.hk>（中文版）或 <http://www.google.com>（英文版）。如图 2-1 所示，Google 主页默认配置即为网页搜索。单击中文主页下方的“Google.com in English”链接可以转入英文主页。



图 2-1 Google 检索入口

Google 网页搜索提供了简单搜索和高级搜索两种检索方式，默认采用简单搜索模式：在输入框中键入一个或多个搜索关键词，然后按下回车键或单击“Google 搜索”按钮即可启动搜索。随后，Google 就会生成搜索结果页面，即与输入的搜索关键词相关的网页列表，搜索结果中的各条记录按照相关度由高到低的次序排列。

单击 Google 主页的输入框右侧的“高级搜索”链接即可进入高级搜索页面。

在 Google 搜索结果页面顶部的“Google 搜索”按钮下方也有“高级搜索”链接，可以从搜索结果页面进入高级搜索页面。高级搜索页面通过提供用户配置入口来设置搜索条件，图 2-2 显示了高级搜索页面，可配置的选项分为以下几部分。

A. 搜索结果

左侧的 4 个输入栏用于配置搜索关键词之间的关系，具体如下：

(1) 包含全部字词：输入的关键词之间是“与”关系；

(2) 包含完整字句：输入的多个关键词构成词组，相当于在搜索输入框用引号包含此栏输入的所有关键词；

(3) 包含至少一个字词：输入的关键词之间是“或”关系，相当于在搜索输入框用 OR 连接此栏输入的多个关键词；

(4) 不包括字词：此栏输入的关键词与其他栏输入的关键词之间是“非”关系，相当于在搜索输入框中在此栏输入的关键词前加上英文的“-”标记。

B. 每页数量

可用于配置每页搜索结果页面显示的结果项数。位于 A 部分右侧的下拉菜单，用于设置搜索结果页面所显示的查询结果的条数；默认为每页 10 条，也可以根据用户的习惯将其设置为每页 20、30、50 或 100 条。

C. 文件格式

限制搜索范围在以特定文件格式创建的网页中，相当于在搜索输入框中使用 filetype：算符作为文件格式的前缀。

D. 日期

将搜索范围限制在特定时间段内更新过的网页中，可选项有：任何时间、过去一天内、过去一周内、过去一个月内、过去一年内，相当于在搜索结果页面设置“百宝箱”的时间选项。

E. 字词位置

限制搜索关键词出现在网页中的位置，除了默认位置是网页内的任何地方外，其他可选项包括：网页的标题、网页的正文、网页的网址、指向网页的链接，相当于分别在搜索输入框中关键词前加上 allintitle:、allintext:、allinurl: 或 allinanchor: 算符。

F. 网站

限制搜索范围在特定的域或网站内，相当于在搜索输入框中使用 site: 算符作为域或网址的前缀。

G. 类似网页

搜索类似网页，相当于在搜索输入框中使用 related: 算符作为特定网址的前缀。