

电子商务员基础知识：常用检索工具（1）电子商务考试 PDF 转换可能丢失图片或格式，建议阅读原文

[https://www.100test.com/kao\\_ti2020/514/2021\\_2022\\_\\_E7\\_94\\_B5\\_E5\\_AD\\_90\\_E5\\_95\\_86\\_E5\\_c40\\_514734.htm](https://www.100test.com/kao_ti2020/514/2021_2022__E7_94_B5_E5_AD_90_E5_95_86_E5_c40_514734.htm) 1.检索工具的分类（1）

按照信息搜集方法分类 按照信息搜集方法的不同，搜索引擎系统可以分为三大类：1）目录式搜索引擎（Directory

Search Engine）以人工方式或半自动方式搜集信息，由编辑员查看信息之后，人工形成信息摘要，并将信息置于事先确定的分类框架中。信息大多面向网站，提供目录浏览服务和直接检索服务。该类搜索引擎因为加入了人的智能，所以信息准确、导航质量高，缺点是需要人工介入（维护工作量大）、信息量少、信息更新不及时。这类搜索引擎的代表是

：Yahoo！、LookSmart、Ask Jeeves、Snap、Open Directory. 2

）机器人搜索引擎（Crawler-Based Search Engine）由一个称为蜘蛛（Spider）的机器人程序以某种策略自动地在Internet中搜集和发现信息，由索引器为搜集到的信息建立索引，由检索器根据用户的查询输入检索索引库，并将查询结果返回给用户。服务方式是面向网页的全文检索服务。该类搜索引擎的优点是信息量大、更新及时、毋需人工干预，缺点是返回信息过多，有很多无关信息，用户必须从结果中筛选。这类搜索引擎的代表是：AltaVista、Northern Light、Excite、Infoseek、Inktomi、FAST、Lycos、Google. 3）元搜索引擎（Meta

Search Engine）这类搜索引擎没有自己的数据，而是将用户的查询请求同时向多个搜索引擎递交，将返回的结果进行重复排除、重新排序等处理后，作为自己的结果返回给用户。服务方式为面向网页的全文检索。这类搜索引擎的优点是返

回结果的信息量大，缺点是不能够充分使用元搜索引擎的功能，用户需要做更多的筛选。这类搜索引擎的代表是WebCrawler、InfoMarket。目前，商业的搜索引擎站点正在结合各种搜索引擎的优点，在类型上有逐渐融合的趋势。例如，Yahoo！在保持人工分类的同时，使用Inktomi的机器人搜索引擎，用户查询时，如果选“网站搜索”便搜索人工分类库，选择“网页搜索”便搜索机器人搜索引擎的索引库。一些传统的机器人搜索引擎也增加了人工分类的内容，以提供高精度的导航信息。另外搜索引擎站点有“门户化”的倾向，在提供搜索服务的同时，提供多样的网络服务，如新闻、股票、天气预报、虚拟社区、游戏、电子商务等等，成为名副其实的“网络门户”。

(2) 按照检索软件分类 按照服务提供方式的不同，检索软件也可以分为三大类：全文数据库检索软件、非全文数据库检索软件、主题指南类检索软件。全文数据库检索软件正常运作的前提是网站拥有大量的信息，因此必须依靠强大的数据库作为后盾。它能够提供完整的文献和信息检索，查全率很高。但由于信息量非常大，检索起来比较困难，对检索技术的要求很高。非全文数据库检索软件具有速度快、使用简便、索引量大的特点，但仅提供部分全文检索，有时需要二次检索，感到不太方便。主题指南类检索软件是目前网络检索中最常用的检索软件。这种软件查准率高、速度快、使用方便。现大部分网站都具备主题指南类检索功能。

(3) 按照检索语言分类 目前，因特网几乎使用了世界所有语言。每一种语言都形成了自己独特的检索体系。比较常用的语言有英文、法文、德文、日文、俄文、中文等。

## 2.全文数据库的检索软件

(1) Alta Vista 搜索引擎 Alta

Vista检索引擎为数字设备公司（DEC）开发，号称是目前最大的Web索引数据库，Alta Vista检索引擎提供两种检索方法：简单检索和高级检索。高级检索包括了简单检索的所有特性，还允许使用布尔运算符和接近操作符、括号等，查找的结果按关键词排序。Alta Vista总能返回有用的信息，但由于没有对内容进行选择，它的“信噪比”也是最大的。使用Alta Vista检索引擎时应注意：要进行有效的检索，最好尽可能多而精确地输入描述所感兴趣的主题的词或词组。提供的词组越精确，检索结果就越好。如果你输入的词包含大写字母，则检索对大小写就比较敏感。如输入Telephone则只检索含有这个词的内容，而输入telephone则不论大小写都检索。如果要把一些词作为词组或一个整体来查询时，最好把它们加上双引号。如果要求特定单词包含在索引的文档中，可以在它前面加一个" "号，如：Telephone，并且在" "号和单词之间不能有空格。相应的，如果要排除含有特定单词的文档，可以在它前面加一个" - "号，如 - cool. 进行简单查找的时候，可以在单词的末尾加一个通配符来代替任意的字母组合（最多可代替5个小写的字母）。Alta Vista的通配符是"\*"号，星号不能用在单词的开始或中间。Alta Vista高级检索包含了简单检索的所有特性，还可以有布尔和接近操作符、括起来的逻辑组合等。Alta Vista支持的布尔和接近检索二元操作符有AND（&）、OR（|）、NEAR（~）和一元操作符NOT（！）。AND连接的若干词在文件中要同时出现；OR 确保检索式中至少有一个词出现在文件中；NOT 将某一个词从检索中排除出去。它们的优先级是递减的，另外要注意，如果把检索的表达式写成sports NOT swimming，则语法上是不合逻辑

的，正确的写法应该是sports AND NOT swimming. NEAR 确保查询的两个词在10个字节内出现，它的优先级是最低的。如果将上述符号用在检索表达式中，最好给检索表达式加上引号以减少检索表达式的混乱。Alta Vista的优点是：在所有的检索软件中，它的功能最全面，查全率和查准率最高，全文标引系统质量可靠；在检索语句上与传统的联机检索语言类似，更容易掌握；系统反应速度快；网页链接可靠。缺点是文件相关性评估的质量一般。

(2) Excite检索引擎 Excite检索引擎是Architext软件公司的产品，该数据库界面友好，用户可以利用关键词、词组和自然语言检索，自然语言检索越详细越好。Excite检索引擎有基本检索和高级检索两种检索方式。其中，Excite中要求的单词和排除的单词的使用方法同Alta Vista一样，使用" "号和" - "号。Excite支持二元操作符AND、OR、AND NOT 和一元操作符 NOT.它也支持用括号来构成逻辑组。缺省的关键词使用的是隐式的OR，即它检索含有指定的任意的单词。Excite的优点是采用了概念检索的技术。概念检索是指在检索文件的过程中，不仅能够检索到含有用户提出的关键词的文件，还能检索到与用户的检索主题密切相关、但并没有包括这些主题词的文件。在所有的检索软件中，它的更新速度最快，且其数据库的规模也最大。另外，它在提供相关信息方面表现颇佳。Excite的缺点是它的相关性排序质量一般，检索结果也显得不尽人意。随着世界华人对因特网商业价值的认同，越来越多的华人开始利用因特网传递商业信息，中文网站内容也不断丰富。面对巨大的市场利益的诱惑，出现了越来越多的中文检索工具，一些原来的英文搜索网站也纷纷推出自己的中文网站。

(3) 非全文数据

库的检索软件 Lycos 是最早出现的检索引擎之一，它是以巡视软件为主的非全文数据库检索工具，由美国卡内基梅隆大学的机器翻译中心开发研制。它利用巡视软件自动搜集网页，对其进行分类，并建立数据库，提供关键词和主题查询（主题查询叫做目录服务）的检索服务。用Lycos查询带符号或数字的词时，最好使用说明性的语言，如用"programming languages C"，而不要用"C"；用"convenience stores 7 ~ 11"，而不要用"7 ~ 11"。Lycos用"|"表示所选择的是与某一主题高度相关的文件，如用"telecommunication TV"选择远程通信中与电视密切相关的文件；排除不必要的文件用"-"，如用"telecommunication-radio"选择远程通信不包括无线电。Lycos用"."表示准确匹配该词；用"\$"表示自动截词符号，查询以该词为词根的所有文件。Lycos的优点在于它的速度快、使用简便、索引量大，它的弱点则是它不支持布尔检索和其他如Alta Vista或Excite能提供的一些高级检索。F8F8" 100Test 下载频道开通，各类考试题目直接下载。详细请访问 [www.100test.com](http://www.100test.com)