



链滴

搜索引擎基本工作原理 -

作者: [lbb4511](#)

原文链接: <https://ld246.com/article/1517018008303>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

了解搜索引擎的工作原理对我们日常搜索应用和网站提交推广都会有很大帮助。

■ 全文搜索引擎

搜索引擎的自动信息搜集功能分两种。一种是定期搜索，即每隔一段时间（比如Google一般是28天，搜索引擎主动派出“蜘蛛”程序，对一定IP地址范围内的互联网站进行检索，一旦发现新的网站，会自动提取网站的信息和网址加入自己的数据库。

另一种是提交网站搜索，即网站拥有者主动向搜索引擎提交网址，它在一定时间内（2天到数月不等）定向向你的网站派出“蜘蛛”程序，扫描你的网站并将有关信息存入数据库，以备用户查询。由于近来搜索引擎索引规则发生了很大变化，主动提交网址并不保证你的网站能进入搜索引擎数据库，因此前最好的办法是多获得一些外部链接，让搜索引擎有更多机会找到你并自动将你的网站收录。

当用户以关键词查找信息时，搜索引擎会在数据库中进行搜寻，如果找到与用户要求内容相符的网站便采用特殊的算法——通常根据网页中关键词的匹配程度，出现的位置/频次，链接质量等——计算各网页的相关度及排名等级，然后根据关联度高低，按顺序将这些网页链接返回给用户。

■ 目录索引

与全文搜索引擎相比，目录索引有许多不同之处。

首先，搜索引擎属于自动网站检索，而目录索引则完全依赖手工操作。用户提交网站后，目录编辑人会亲自浏览你的网站，然后根据一套自定的评判标准甚至编辑人员的主观印象，决定是否接纳你的网。

其次，搜索引擎收录网站时，只要网站本身没有违反有关的规则，一般都能登录成功。而目录索引对站的要求则高得多，有时即使登录多次也不一定成功。尤其象Yahoo!这样的超级索引，登录更是困难（由于登录Yahoo!的难度最大，而它又是商家网络营销必争之地，所以我们会在后面用专门的篇幅介绍登录Yahoo雅虎的技巧）

此外，在登录搜索引擎时，我们一般不用考虑网站的分类问题，而登录目录索引时则必须将网站放在一个最合适的目录（Directory）。

最后，搜索引擎中各网站的有关信息都是从用户网页中自动提取的，所以用户的角度看，我们拥有更大的自主权；而目录索引则要求必须手工另外填写网站信息，而且还有各种各样的限制。更有甚者，如工作人员认为你提交网站的目录、网站信息不合适，他可以随时对其进行调整，当然事先是不会和你量的。

目录索引，顾名思义就是将网站分门别类地存放在相应的目录中，因此用户在查询信息时，可选择关键词搜索，也可按分类目录逐层查找。如以关键词搜索，返回的结果跟搜索引擎一样，也是根据信息程度排列网站，只不过其中人为因素要多一些。如果按分层目录查找，某一目录中网站的排名则是由题字母的先后顺序决定（也有例外）。

目前，搜索引擎与目录索引有相互融合渗透的趋势。原来一些纯粹的全文搜索引擎现在也提供目录搜，如Google就借用Open Directory目录提供分类查询。而象Yahoo!这些老牌目录索引则通过与Google等搜索引擎合作扩大搜索范围（注）。在默认搜索模式下，一些目录类搜索引擎首先返回的是自目录中匹配的网站，如国内搜狐、新浪、网易等；而另外一些则默认的是网页搜索，如Yahoo。