

行业专家倾力力作！国内外知名案例解析！

藏锋者 郝永清 编著

SEO 实战指导

知识
策略
案例

教你成为网络推广大赢家！

SEO是一种思想！

★ Google AdWords与百度推广
新浪财经频道的交换链接
wikipedia.org和百度百科

营销

最实用SEO策略指导：

房地产企业的关键词选择
新浪财经频道的交换链接
百度百科的日访问量示例分析
搜搜问问、天涯问答中的外部链接建设

技术操作也没那么难：

百度统计案例
百度指数分析工具案例
友情链接检测工具使用案例
Google Analysis (分析) 流量统计案例

网络
推广

e-Marketing

搜索引擎优化必读！

—实用—

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

藏锋者 郝永清 编著

SEO 实战指导

知识
策略
案例



网络
推广

e-Marketing
搜索引擎优化必读!

—实用—

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书是一本以搜索引擎优化从业者实际工作经验为基础,将外表烦琐,实则规律可循的 SEO 技术内在本质进行提炼的案例分析书籍。本书将搜索引擎优化技术分成优化工作开始前的关键词准备工作、本地站内优化和网站发布后的站外优化推广三个部分,将复杂的 SEO 技术进行提炼和梳理。从作者自身工作要求中进行相关的技术实施,并辅助大量实际操作案例,力争让没有任何基础的读者也能一次上手。

本书适合负责网络营销和网络品牌推广的人员和相关技术人员阅读;适合作为网络营销相关培训、课程的辅助教材阅读。另外,特别推荐想利用网络创造额外收入的上班族、有网络创业梦想的人阅读。

图书在版编目(CIP)数据

SEO 实战指导:知识·策略·案例/藏锋者,郝永清编著. — 北京:中国铁道出版社,2011.5
ISBN 978-7-113-12689-6

I. ①S… II. ①藏… ②郝… 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 036318 号

书 名: SEO 实战指导:知识·策略·案例
作 者: 藏锋者 郝永清 编著

责任编辑: 苏 茜
编辑助理: 张 丹
封面设计: 张 丽
责任印制: 李 佳

读者热线电话: 400-668-0820

出版发行: 中国铁道出版社(北京市宣武区右安门西街 8 号 邮政编码: 100054)

印 刷: 北京鑫正大印刷有限公司

版 次: 2011 年 5 月第 1 版 2011 年 5 月第 1 次印刷

开 本: 787mm×1092mm 1/16 印张: 19.25 字数: 443 千

印 数: 3 500 册

书 号: ISBN 978-7-113-12689-6

定 价: 45.00 元

版权所有 侵权必究

凡购买铁道版图书,如有印制质量问题,请与本社发行部联系调换。

前言

搜索引擎优化

Search Engine Optimization

本书作者根据多年的工作经验，将纷繁复杂的搜索引擎优化技术（SEO）进行提炼和梳理，从理论技术介绍到实际案例操作，进行细致详尽的讲解和分析，力求让没有任何基础的读者也能一次上手。

本书特点

本书采用案例引导加实际操作详细的形式，抛弃繁杂的技术理论讲解，除对关键理论进行简明扼要的阐述以外，绝大多数内容是基于实际案例的分析和操作指导，让读者读起来轻松快捷，极大地提高了读者的阅读兴趣。

本书适合读者

本书适合以下人员阅读：

- 各个行业负责网络营销和网络品牌推广的人员及技术人员。
- 网站建设者、网站站长和网站所有者。
- 想利用网络创造额外收入的上班族。
- 喜欢搜索引擎优化技术的爱好者。
- 有网络创业梦想的所有人。

本书脉络结构

本书从内容结构上可分为三篇：

篇 名	内 容
第 1 篇 搜索引擎优化(SEO) 前置知识（包括第 1~3 章）	搜索引擎优化（SEO）前置知识，主要为刚接触搜索引擎优化的新手讲解一些 SEO 的相关知识、搜索引擎喜欢的网站特点和搜索引擎优化的常用工具等
第 2 篇 网站发布前的内部 优化（包括第 4~11 章）	网站发布前的内部优化，主要介绍站内优化的内容，包括关键词、域名、主机空间、CMS、网站结构、内部链接、内容规划和单网页优化。这部分内容采用案例分析的方式，针对网站发布前所需要进行的搜索引擎优化要点进行详细讲解
第 3 篇 网站推出后的优化 与推广策略（包括第 12~16 章）	网站推出后的外部优化与推广，主要介绍网站发布后的站外优化，包括交换链接策略、博客与论坛外部链接策略、百科与分类目录推广策略和 SNS 社区与问答系统流量吸引策略

本书写作团队

本书由藏锋者、郝永清编著，藏锋者网络创业网（<http://V.cangfengzhe.com>）提供技术支持和免费读者服务平台。

由于时间仓促，加之水平有限，书中出现不足之处在所难免，敬请读者批评指正。如果读者在阅读本书的时候出现任何疑问，还可以发送电子邮件或者登录藏锋者网络创业网及时与我们联系，我们会尽快给予答复。

技术支持邮箱：wwb_beijing@163.com

免费创业咨询交流：<http://v.cangfengzhe.com/wenba/>

目 录

搜索引擎优化

Search Engine Optimization

第 1 篇 搜索引擎优化 (SEO) 前置知识

第 1 章 初识搜索引擎优化 (SEO)	2
1.1 什么是搜索引擎优化 (SEO)	2
1.1.1 搜索引擎工作原理和步骤	2
1.1.2 搜索引擎优化 (SEO) 的含义	4
1.1.3 SEO 是最佳方法	5
1.2 搜索引擎优化的目标	5
1.2.1 更好的搜索排名	5
1.2.2 提升用户体验	6
1.2.3 创收与创业	6
1.3 搜索引擎优化的分类	8
1.3.1 站内搜索引擎优化	8
1.3.2 站外搜索引擎优化	8
1.4 搜索引擎优化的误区	9
1.4.1 为了搜索引擎而优化	9
1.4.2 听信虚假广告	9
1.4.3 过于急迫	10
1.4.4 采用黑帽手法	11
1.5 搜索引擎优化的适用范围和人员要求	12
1.5.1 SEO 适合所有有网络创业梦想的人	12
1.5.2 SEO 人员应该具备的素养	12
1.6 小结与案例汇总	13
第 2 章 搜索引擎喜欢的网站特点	14
2.1 主题明确的网站	14
2.1.1 什么是网站主题	14
2.1.2 22 个经典网站主题	14
2.1.3 如何定位网站主题	30
2.1.4 建设与目标主题密切相关的网站	32

2.2	权威的网站	32
2.2.1	什么是网站权威性	32
2.2.2	影响网站权威性的因素	32
2.3	实用的网站	33
2.3.1	网站实用性与用户行为	33
2.3.2	影响网站实用性的用户行为	34
2.4	小结与案例汇总	37
第 3 章	搜索引擎优化 (SEO) 常用工具	38
3.1	网站信息查询工具	38
3.1.1	Alexa 排名查询	38
3.1.2	Google PR 查询和全站 PR 查询	41
3.1.3	反向链接查询	42
3.1.4	每小时、每天、每周、每月搜索引擎收录查询	43
3.1.5	死链接检测	45
3.1.6	网速测试之超级 Ping	45
3.1.7	域名查询、域名删除列表	46
3.1.8	域名 Whois 查询	47
3.1.9	同 IP 网站查询	48
3.1.10	PR 输出值查询	49
3.2	流量统计系统	50
3.2.1	Google Analytics 的应用	50
3.2.2	百度统计的应用	60
3.2.3	量子统计的应用	65
3.3	搜索优化工具	71
3.3.1	百度指数分析	72
3.3.2	Google 相关关键词查询	72
3.3.3	关键词排名查询	73
3.3.4	关键词密度查询	74
3.3.5	友情链接检测	75
3.3.6	搜索引擎蜘蛛模拟	75
3.4	搜索引擎官方推荐工具	76
3.4.1	搜索引擎优化必备工具: Google 网站管理员工具	76
3.4.2	百度站长工具	84
3.5	小结与案例汇总	85

第 2 篇 网站发布前的内部优化

第 4 章	SEO 第一步: 关键词策略	88
--------------	-----------------------------	-----------

4.1	初识关键词	88
-----	-------------	----

4.1.1	什么是关键词	88
4.1.2	关键词与搜索密切相关	88
4.1.3	关键词确定	89
4.2	关键词选择原则	90
4.2.1	关键词与网站主题要紧密相关	90
4.2.2	关键词要精准而不能宽泛	90
4.2.3	关键词不要太特殊	92
4.2.4	站在用户角度思考	92
4.2.5	选择竞争度最小的关键词	93
4.3	关键词选择步骤实例分析	93
4.3.1	建立简易关键词库	93
4.3.2	关键词竞争度评估	96
4.3.3	长尾关键词策略	99
4.4	小结与案例汇总	101
第 5 章 域名与 SEO		102
5.1	域名权重	102
5.1.1	Google 专利中的域名权重	102
5.1.2	域名权重的常规意义	103
5.1.3	域名权重分级	104
5.1.4	域名权重的争议和逻辑推理	104
5.1.5	Google、百度的另一身份: 域名注册商	105
5.2	域名选取的原则	106
5.2.1	适合记忆	106
5.2.2	输入方便	108
5.2.3	有丰富内涵	109
5.3	域名选取的技巧及实例	109
5.3.1	用主关键词作为域名	109
5.3.2	用产品作为域名	110
5.3.3	用品牌作为域名	110
5.3.4	用企业名作为域名	111
5.3.5	SEO 中域名选择的误区	113
5.4	案例: 如何打造适合搜索引擎的域名	114
5.4.1	要点一: 方便记忆并包含关键词	115
5.4.2	要点二: 权重链接指向	116
5.4.3	要点三: 域名注册时间和历史	122
5.4.4	要点四: 域名所有人的变更	123
5.4.5	要点五: 域名下内容的质量	124
5.4.6	要点六: 域名中信息的更新	125
5.4.7	要点七: 域名下包含的页面数量	126

5.4.8	要点八: 高质量的出站链接	127
5.4.9	要点九: 死链接与 404 页面的处理	127
5.4.10	要点十: 稳定的服务器	128
5.4.11	要点十一: 域名中页面的标准	128
5.5	小结与案例汇总	129
第 6 章	主机空间与 SEO	130
6.1	主机和空间基础知识	130
6.1.1	什么是主机, 什么是空间	130
6.1.2	购买独立主机与网络空间的抉择	130
6.2	主机空间对 SEO 的影响	131
6.2.1	主机空间速度对 SEO 的影响	131
6.2.2	主机空间稳定性对 SEO 的影响	132
6.2.3	主机设置错误对 SEO 的影响	132
6.2.4	主机空间的地理位置对 SEO 的影响	133
6.2.5	同一个 IP 上的其他网站对 SEO 的影响	133
6.3	选择优秀主机空间的方法	134
6.3.1	选择稳定主机空间的方法	134
6.3.2	选择高速主机空间的方法	135
6.3.3	选择更多连接数的主机空间	136
6.3.4	尽可能地购买更多的流量	137
6.3.5	选择有备份机制的主机空间	137
6.3.6	自定义 404 页面	138
6.3.7	子域名支持	138
6.3.8	服务响应时间	139
6.4	国内各类型主机空间推荐	140
6.4.1	主机托管推荐	140
6.4.2	万网虚拟主机推荐	142
6.4.3	VPS 主机推荐	142
6.5	小结与案例汇总	143
第 7 章	适合 SEO 的 CMS 程序选择与调试	144
7.1	CMS 系统简介	144
7.1.1	什么是 CMS	144
7.1.2	CMS 的分类	145
7.2	CMS 在搜索引擎优化中的优势	145
7.2.1	CMS 的 HTML 生成功能有利于规范网站 URL 结构	146
7.2.2	CMS 利于灵活定义网页标题、关键词和描述标签	146
7.2.3	CMS 有助于优化导航和网站地图	147
7.3	国内适合 SEO、SEM 的主流 CMS 概览	148

7.3.1	PHPCMS.....	148
7.3.2	DedeCMS	152
7.3.3	PHP168.....	154
7.3.4	动易 CMS.....	155
7.4	用 CMS 建立本地调试网站	158
7.4.1	下载运行环境 (pc_webserver)	158
7.4.2	安装运行环境 (pc_webserver)	159
7.4.3	访问本地网站	159
7.4.4	登录后台生成网站基础框架	160
7.5	小结与案例汇总	162
第 8 章	确立搜索引擎喜欢的网站结构.....	163
8.1	网站结构介绍	163
8.1.1	什么是网站结构	163
8.1.2	物理结构与逻辑结构	163
8.1.3	网站结构对 SEO 的影响.....	165
8.2	网站结构优化	165
8.2.1	物理结构优化	165
8.2.2	逻辑结构优化	166
8.3	二级域名与子目录的选择.....	167
8.3.1	根据内容的丰富程度进行选择.....	167
8.3.2	根据内容的相关性进行选择	167
8.3.3	分开访问子域名内容	168
8.3.4	不要滥用子域名	168
8.4	小结与案例汇总	168
第 9 章	内部链接优化.....	169
9.1	初识链接	169
9.1.1	什么是链接	169
9.1.2	链接分类	170
9.2	内部链接在 SEO 中的作用	171
9.2.1	内部链接有助于加快网站收录.....	172
9.2.2	内部链接有助于优化页面排名.....	172
9.2.3	内部链接有助于加强权重传递.....	173
9.2.4	内部链接有助于提高用户体验.....	173
9.3	内部链接检测工具.....	173
9.3.1	Google 网站管理员工具中的内部链接统计	173
9.3.2	内部链接检查器	175
9.4	新浪网的内部链接规划案例.....	177
9.4.1	扁平的树形网状结构	177

9.4.2	网站导航的规划与部署	182
9.4.3	用链接相关性提高用户体验	186
9.4.4	死链接与错误页面	186
9.5	小结与案例汇总	187
第 10 章	“内容为王”与 SEO 中的内容策略	188
10.1	原创、转载与伪原创	188
10.1.1	什么是原创内容	188
10.1.2	什么是转载内容	189
10.1.3	什么是伪原创内容	190
10.1.4	搜索引擎优化者的选择	192
10.2	内容在 SEO 中的重要作用	193
10.2.1	有价值的内容是搜索引擎优化的第一位	193
10.2.2	以内容为本营造权威品牌	194
10.2.3	SEO 和内容建设相辅相成	194
10.3	标题策略	194
10.3.1	中规中矩的常规标题	195
10.3.2	明确受众的特定标题	195
10.3.3	用数字说话的标题	196
10.3.4	用疑问引起共鸣的标题	196
10.3.5	使用夸张的标题	196
10.3.6	用明星效应武装标题	197
10.3.7	限时的标题	197
10.3.8	不符合常理的标题	197
10.3.9	用热点事件伪装标题	198
10.3.10	抓住访问者好奇心的标题	198
10.4	内容形式	198
10.4.1	不要用太正式、古板的语言	198
10.4.2	让内容建设更简洁	199
10.4.3	多做平实、浅显易懂的内容	199
10.5	内容结构	199
10.5.1	不要写太长的段落	200
10.5.2	多用小标题	200
10.5.3	多罗列要点	201
10.6	表现形式	202
10.6.1	数字的说服力	202
10.6.2	善用富媒体	202
10.7	小结与案例汇总	203

第 11 章 单网页优化	204
11.1 单网页的分类与内容构成	204
11.1.1 首页的内容构成	204
11.1.2 栏目页的内容构成	205
11.1.3 列表页的构成	207
11.1.4 内容页的构成	208
11.1.5 其他单网页的内容构成	209
11.2 兼顾 SEO 与用户体验的单网页优化	209
11.2.1 导航与底部区域	209
11.2.2 内容区域	211
11.2.3 相关内容	212
11.2.4 互动机制	213
11.2.5 代码优化	213
11.2.6 URL 优化	220
11.2.7 关键词布局与密度	222
11.2.8 导入与导出链接	223
11.3 小结与案例汇总	224

第 3 篇 网站推出后的优化与推广策略

第 12 章 外部链接与 SEO	226
12.1 外部链接在 SEO 中的作用	226
12.2 判断优秀外部链接的标准	227
12.3 让别人主动链接：链接诱饵	230
12.3.1 什么是链接诱饵	230
12.3.2 链接诱饵的常用手法	230
12.4 小结与案例汇总	232
第 13 章 交换链接策略	233
13.1 交换链接基础知识	233
13.1.1 什么是交换链接	233
13.1.2 交换链接的意义	233
13.2 优秀交换链接的标准	234
13.2.1 网站相关性	234
13.2.2 网站质量	235
13.2.3 导出链接数量	237
13.2.4 链接伙伴的健康度	238
13.2.5 交换链接的稳定性	238

13.3 寻找优秀交换链接的方法.....	239
13.3.1 搜索引擎.....	239
13.3.2 分类目录与网址导航.....	240
13.3.3 链接交换平台.....	242
13.3.4 同类网站的交换链接.....	243
13.3.5 站长群与站长论坛.....	244
13.4 小结与案例汇总.....	245
第 14 章 构建高质量的博客和论坛外部链接策略.....	246
14.1 博客与论坛外部链接建设的基础知识.....	246
14.1.1 什么是博客外部链接.....	246
14.1.2 什么是论坛外部链接.....	247
14.1.3 博客和论坛外部链接的意义：让外部链接可控.....	248
14.2 博客和论坛外部链接建设的五大误区.....	249
14.2.1 杜绝群发机器人.....	249
14.2.2 数量与质量的博弈.....	250
14.2.3 培养博客、论坛的优秀账户.....	250
14.2.4 原创与转载的融合.....	251
14.2.5 外部链接方式的多样化.....	251
14.3 利用博客、论坛建设高质量外部链接的方法.....	253
14.3.1 建设高质量博客进行外部链接的典型步骤.....	253
14.3.2 论坛外部链接建设三要素.....	257
14.4 小结与案例汇总.....	259
第 15 章 百科与分类目录外部链接建设策略.....	260
15.1 基础知识.....	260
15.1.1 认识百科.....	260
15.1.2 浅谈分类目录.....	264
15.1.3 百科、分类目录外部链接建设的价值.....	265
15.2 百科外部链接建设案例.....	266
15.2.1 国外维基百科外部链接建设案例.....	266
15.2.2 百度百科外部链接建设和推广案例.....	269
15.3 分类目录外部链接建设案例.....	274
15.3.1 登录 hao123.....	274
15.3.2 domz 提交与登录.....	276
15.4 小结与案例汇总.....	278
第 16 章 SNS 社区和问答系统的流量吸引策略.....	279
16.1 SNS 社区的流量吸引策略.....	279

16.1.1	新兴的 SNS 社区	279
16.1.2	最呆板的 SNS 流量吸引方法	280
16.1.3	利用 SNS 进行病毒式的流量吸引	281
16.1.4	常见 SNS 推广代码	281
16.2	问答系统的流量吸引策略	283
16.2.1	问答系统	283
16.2.2	问答系统中的流量吸引方法	285
16.3	小结与案例汇总	288
附录 A	本书案例速查	289

第 1 篇 搜索引擎优化(SEO)前置知识

搜索引擎优化

Search Engine Optimization

在开始进行搜索引擎优化的各项具体操作之前，针对搜索引擎优化新手，本篇将首先介绍搜索引擎优化的基本概念和常规注意事项，目的是为了让新手了解 SEO 的定义、范围和目标。

第1章 初识搜索引擎优化 (SEO)

搜索引擎优化

Search Engine Optimization

本章介绍:

- ✎ 搜索引擎优化的定义
- ✎ 搜索引擎工作原理
- ✎ 搜索引擎优化的目标、分类和误区
- ✎ 搜索引擎优化人员的基本素养
- ✎ 百度与 Google 蜘蛛 (Spider) 的日志记录
- ✎ 百度拼写错误处理
- ✎ Google 网站管理员工具中的外部链接统计
- ✎ 伪原创与博客群发工具

搜索引擎优化是目前被广为关注的网站建设、网络推广手段,无数网站建设者、网络营销人员都已经参与到其中,并且也有很多成功的案例逐渐公布出来。作为搜索引擎优化新手,在开始进入搜索引擎优化的具体技术学习之前,应当先充分了解搜索引擎优化的基本定义、方向、优势和误区等内容,这对以后的实际操作有评估价值和指导意义。

1.1 什么是搜索引擎优化 (SEO)

搜索引擎优化由“搜索引擎”和“优化”两个部分组成,其中“搜索引擎”是平台,“优化”是动作。

1.1.1 搜索引擎工作原理和步骤

要了解搜索引擎优化,首先要了解搜索引擎的基本工作原理。从这个原理出发,才可以摸索出搜索引擎优化更深层次的内涵。

 **提示:**“搜索引擎”分类很繁杂,可以包含全文索引、目录索引、元搜索、垂直搜索等。本书中提到的“搜索引擎”,特指全文索引的搜索引擎,也就是日常在网络中使用的谷歌、百度等主流搜索引擎平台。

通常情况下,搜索引擎的工作过程分为以下四个步骤。

1. 爬行和抓取

搜索引擎开始工作的时候,首先会派出一个能够在网上发现新网页并抓取文件的程序(这个程序通常被称为蜘蛛 Spider、爬虫 Crawler 或机器人 Robot),这个程序从搜索引擎自身数据库中已知的网页出发,像正常用户的浏览器一样访问已存在的网页上的链接,并根据链接抓取文件。

 **提示：**“蜘蛛”(Spider)、“爬虫”(Crawler)、“机器人”(Robot)这三种不同叫法的实质意义其实类似，在本书中不做严格区分。一般情况下，可以理解为三者可以相互替换，如图1-1和图1-2所示。

```
61.135.168.82 - - [18/Apr/2009:01:09:16 -0700] "GET / HTTP/1.1" 200 13032 "-" "Baidu
220.181.32.22 - - [18/Apr/2009:01:12:46 -0700] "GET / HTTP/1.1" 200 13032 "-" "Baidu
61.135.190.231 - - [18/Apr/2009:01:45:19 -0700] "GET / HTTP/1.1" 200 7491 "-" "Baidu
61.135.190.240 - - [18/Apr/2009:01:48:24 -0700] "GET /product/index-1.htm HTTP/1.1"
220.181.32.22 - - [18/Apr/2009:01:51:27 -0700] "GET /hua/index-10.htm HTTP/1.1" 200
61.135.190.244 - - [18/Apr/2009:01:54:30 -0700] "GET /hua/1330.htm HTTP/1.1" 200 254
61.135.190.239 - - [18/Apr/2009:01:57:33 -0700] "GET /hua/1329.htm HTTP/1.1" 200 256
61.135.190.225 - - [18/Apr/2009:02:00:36 -0700] "GET /hua/1328.htm HTTP/1.1" 200 256
61.135.190.232 - - [18/Apr/2009:02:02:50 -0700] "GET /hua/1327.htm HTTP/1.1" 200 256
```

图1-1 百度蜘蛛分析

```
2009-03-04 00:24:53 W3SVC/1.210.53.60 GET /bbspace.php username=wyx559.00 - 200 208 60 45
Mozilla/5.0+(compatible;+Googlebot/2.1.+http://www.google.com/bot.html) 301 0 0
2009-06-04 00:24:35 W3SVC/303059524.210.85.80 6 GET /space.php username=wyx559.80 - 200 208 60 37
Mozilla/5.0+(compatible;+Googlebot/2.1.+http://www.google.com/bot.html) 200 0 0
2009-04-03 00:13:06 W3SVC/1.210.63.80 6 GET /bbs/viesthread.php fid=607&page=1&authorid=1007.50 -
303 206 60 45 Mozilla/5.0+(compatible;+Googlebot/2.1.+http://www.google.com/bot.html) 301 0 0
2009-09-03 00:13:06 W3SVC/193060524.210.85.80 6 GET /weathread.php fid=607&page=1&authorid=1007.80 -
200 208 60 37 Mozilla/5.0+(compatible;+Googlebot/2.1.+http://www.google.com/bot.html) 200 0 0
```

图1-2 Google 蜘蛛

“蜘蛛”在访问已知的网页后，会跟踪网页上的链接，并访问更多的网页，这个过程叫做爬行。当通过链接发现有新的网址时，“蜘蛛”就把新的网址记入搜索引擎自己的数据库，等待抓取。

跟踪网页链接是搜索引擎“蜘蛛”发现新网址的最基本方法，所以网页链接的优化也就成为搜索引擎优化的最基本方法之一。

搜索引擎“蜘蛛”抓取的页面文件，往往与用户浏览器中看到的页面大不相同，“蜘蛛”会将这些抓取的网页文件存入数据库，以待后用。

2. 索引

搜索引擎程序把“蜘蛛”抓取的网页文件分解、分析，并以某种特定的形式存入自己的庞大数据库，这个过程就是索引。

在索引数据库中，网页的文字内容、关键词出现的位置、字体、颜色等信息都有相应的记录。

 **提示：**搜索引擎索引数据库存储巨量数据，主流搜索引擎通常都有几亿、几十亿条网页。

3. 搜索词处理

用户在搜索引擎界面输入关键词，单击“搜索”按钮后，搜索引擎程序立即对输入的搜索词进行处理，如图1-3所示。

这个处理过程很烦琐，而且中间的过程对用户而言是不可见的，也是搜索引擎的核心机密之一。常见的搜索词处理包括中文的分词、关键词词序的辨别、去除停止词、判断是否需要启动整合搜索、判断是否有拼写错误或错别字等，如图1-4所示。



图1-3 搜索词处理

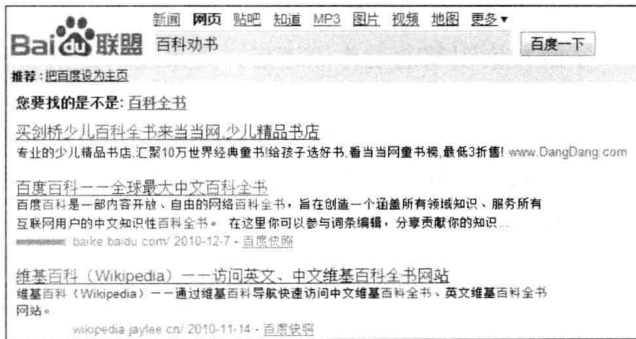


图1-4 搜索引擎的拼写错误处理