

网站 SEO 实证研究

柴庆慧 贾素玲

(北京航空航天大学,北京 100083)

E-mail:qinghuichai@163.com

摘要 针对搜索引擎优化(SEO)在网站推广中的重要性,提出了现实中搜索引擎优化的部分方法,并对其进行了实证研究。

关键词 搜索引擎优化(SEO) 网站推广

Research on the SEO of Website

Chai Qinghui Jia Suling

(Beijing University of Aeronautics & Astronautics, Beijing 100083)

Abstract: Owing to the importance of SEO in website popularization, proposes some methods in the realization of SEO, and does demonstration research on it.

Keywords: Search Engine Optimization, website popularization

1 引言

在网站开发实践中,多数人比较注重采用先进的设计模式,使得网站的扩展性、功能完备性、美工设计等方面都比较满意,但对网站推广这关键的一环却往往忽略了。在互联网高速发展、信息量急剧扩大的今天,搜索引擎的作用日益突出,企业或组织都想借助网络平台让更多人了解自己或进行商务活动,因此搜索引擎优化(SEO; Search Engine Optimization)显得尤其重要。本文在对一家房地产顾问公司网站 www.rl-consult.com 改版的策划、建设、维护的全过程进行追踪的基础上,对搜索引擎优化工作进行了总结和研究。

2 域名和主机对 SEO 的影响

域名与主机是网站的基础,商业网站选择域名与主机尤其注重形象和质量。同时选择域名与主机也是搜索引擎优化的第一步。由于是对 www.rl-consult.com 进行改版,该公司已经使用此域名四年多了,从营销角度讲,更换域名是一种不理智的行为,故没有进行改变。以前旧版网站挂在 263 的虚拟主机上,调研后发现虚拟主机有很多不利于 SEO 进行的问题:

(1) 虚拟主机上的网站如果被搜索引擎惩罚,将或多或少波及自己的网站; AltaVista 有如下忠告:“如果同一 IP 下一个网站作弊,那么我们的搜索引擎会对该 IP 下的所有网站进行惩罚”;幸运的是,大部分搜索引擎不会如此武断,如 Google 一般不进行这种连带惩罚。

(2) 如果搜索引擎对有些 Spammer 服务器进行惩罚,其 IP 被认为是 Spam,则该 IP 下的所有网站都会受到牵连,使这些网站的排名下降或在搜索引擎中被清除。

(3) 虚拟主机会对每个网站流量进行限制,有了流量制约,一个网站很难取得成功。

针对以上问题,网站建设者决定购买自己的服务器,进行主机托管,考虑到带宽也会对一个网站产生至关重要的影响,所以选择服务器托管商也需要下一番工夫,经过现场考察与综

合比较,最终选定中关村机房,共享 100M 带宽。

3 动态网页静态化

所谓动态网页一般是指采用 ASP、PHP、Cold Fusion、CGI 等程序动态生成的页面,该网页中大部分内容来自与网站相连的数据库。由于访问者能实时得到他们想要的数据库,动态网页往往给人留下深刻的印象。此外,动态网页还具有易维护和更新的优点。但大多数搜索引擎的蜘蛛程序都无法解读符号“?”后的字符。这就意味着动态网页很难被搜索引擎的 Spider 检索到,因而被用户找到的机会也大为降低。因此,网站运行时最好使用技术将动态网页转化成静态网页,使 URL 中不再包含“?”“=”等类似的符号。也可以对网站进行改动间接增加动态网页的搜索引擎可见度。即坚持“动静结合,以静制动”的原则。

基于以上原因,又考虑到旧版 www.rl-consult.com 是用 PHP 编写的动态网站,因此改版后编程时第一件事就是进行页面的静态化。

4 对搜索引擎友好的网页设计制作原则

一般的网页都由网页设计师完成。设计师设计网站往往仅从美观、创意和易用的角度考虑,这对于期望获得搜索引擎排名优秀的商业网站来说是远远不够的,目前 SEO 的出现,给网页设计师们提出了一个新的挑战。

4.1 目录层次

小型网站一般只有一层子目录,例如: http://www.rl-consult.com/danye/zhaopin.htm, 其中 www.rl-consult.com 是域名, danye 是一级目录名, zhaopin.htm 是文件名。对搜索引擎而言,这种目录结构最为理想,即扁平结构(Flat)。而规模大一些的网站往往需要更多层子目录,如: http://www.rl-consult.com/html/ziliao/ziliao2/。搜索引擎还是会抓取二到三层子目录下的文件,但最好不要超过三层,四层以上就很难搜索了。出现四层结构,也可以 SEO,像 www.rl-consult.com 因为内容的需要,将目录

作者简介:柴庆慧(1982-),女,硕士研究生,研究领域为管理信息系统。贾素玲,女,北京航空航天大学教授,主持过多项国家自然科学基金项目。

层次设计到第四层,然后又用一种特殊的处理方法:在首页上增加一个该页链接,可以通过首页直接到达。最终在 Google 中检索该页时,竟然能够检索到,说明这种优化方法还是比较成功的。

4.2 使用网站地图

网站地图(Site Map)是辅助导航的手段,最初是为用户设计的,以方便用户快捷到达目标页。网站地图常常用网站拓扑结构体现复杂的目录关系,具有静态、直观、扁平、简单的特点。多采用文本链接,不用或少用修饰性图片,以加快页面加载速度。网站地图在 SEO 中有重要的意义。对于那些采用图片导航和动态技术生成的网页以及较深层次的目录,增加文本链接,可在一定程度上弥补蜘蛛程序无法识别图片、动态网页、深层目录不可见的风险。需要注意,网站地图也要突出重点,尽量给出主干性内容及链接,而不是所有细枝末节。一页内不适宜放太多链接。Google 明确提出“如果网站地图上的链接超过大约 100 个,则最好将网站地图拆成多个网页”。

针对以上问题,网站建设者专门做了一个网站地图页面,将 www.rl-consult.com 中重要页面、深层目录做了文字链接,并将其在主页做了链接。

4.3 图像优化

一般,搜索引擎只识读文本内容,对图像是不可见的。同时,图像文件延缓页面加载,如果超过 20 秒还不能加载,用户和搜索引擎极有可能离开。因此,除非网站内容以图片为主或者图片至关重要,否则尽量避免使用大图片,更不要采用纯图像制作网页。

网站图片优化的核心有两点:增加搜索引擎可见的文本描述,及在保持图像质量的情况下尽量压缩图像文件。增加文本描述,即标签中的 alt 属性,Google、baidu 对这一内容比较重视。在 www.rl-consult.com 首页中每个图片都添加了此属性,并带上整个网站的关键字,如:“房地产顾问 房地产咨询 房地产策划”。注意,中间用空格或逗号分隔。

4.4 关键字密度

这可能是 SEO 中谈论最多的话题。人们喜欢用一个页面中的关键字密度来作为搜索引擎排名的一个参考因素而不是出现关键字或关键字短语的次数。所谓关键字密度是指关键字与一个页面中除掉 html 代码后内容的百分比。研究表明,大多数搜索引擎中关键字密度 2%~8% 比较适当,有利于网站在搜索引擎中排名。但经过观察,这个范围不是绝对的,应根据内容确定关键字密度。实证中发现 www.rl-consult.com 主页中关键词密度超过 8%,仍有较高的排名。如“地产 顾问”在各大引擎的排名如表 1,其中,TOM 网页搜索内核为 yahoo.cn,Sohu 搜索内核为 Sogou,网易搜索内核为 Google。统计时间为 2006-03-09。

表 1 搜索引擎排名

搜索引擎	排名
Yahoo.cn	3
Google	6
Baidu	11
Sogou	6

一般,优化的不是一个词语而是一个短语,如“房地产顾问”就包括了“房地产”与“顾问”两个词语,优化时不但要强调“房地产顾问”还要强调“房地产”与“顾问”。(注:所优化的“房

地产顾问”是一个短语而不是一个词语),搜索引擎在识别关键字短语时会把短语整个解释及分开成多个词语解释。为提高排名,在 www.rl-consult.com 首页的左侧与底部均加入了关键词,在底部,以版权形式加入关键词,不显突兀,祛除作弊之嫌。

4.5 <meta>标签

在网页中放置 META 标签,如果不是滥用的话,对于网站排名是很好的。keyword 标签仅仅包含页面主要的关键字而无需在页面中多次重复。至于 description 标签,一般在 25-30 或者更少一些为好。有些搜索引擎会利用这个作为网站的描述。如 Google 就有时利用它作为网站的描述。如: http://www.dll-sz.com/ keywords,description 中都是这样写的:“深圳,房地产,中介,天驷行,天驷,顾问咨询,企业咨询,DLL,代理,市场研究,营销,策划,销售策划,销售代理,广告咨询,投资代理,设计咨询,前期策划,房地产评估”。Google 对此网站的描述为:

深圳市天驷行房地产顾问有限公司,深圳,房地产,中介,天驷行,天驷,顾问...
 深圳,房地产,中介,天驷行,天驷,顾问咨询,企业咨询,DLL,代理,市场研究,营销,策划,销售...
 代理,广告咨询,投资代理,设计咨询,前期策划,房地产评估
 www.dll-sz.com/~16k- 网页快照 - 类似网页

5 Google, Alexa 实战解析

www.rl-consult.com 中三组关键字“房地产 策划”、“房地产 咨询”、“房地产 顾问”在 Google 中的波动情况,如表 2 (2006-02-06~2006-03-09)。

表 2 在 Google 中的波动情况

日期	房地产 咨询	房地产 策划	房地产 顾问
2006-02-06	800	800	10
2006-02-08	137	500	8
2006-02-12	500	129	6
2006-02-13	142	126	5
2006-02-15	139	129	5
2006-02-18	101	101	6
2006-02-20	97	96	4
2006-02-22	89	81	5
2006-02-24	82	91	4
2006-03-09	87	89	5

www.rl-consult.com Alexa 世界排名波动情况如表 3 所示 (2005-12-09~2006-03-09)。

表 3 www.rl-consult.com Alexa 世界排名波动情况

Traffic Rank for rl-consult.com (what's this)			
Today	1 wk. Avg.	3 mos. Avg.	3 mos. Change
-	862,661	770,931	◆ 794,820
Reach for rl-consult.com (what's this)			
Reach per million users:			
Today	1 wk. Avg.	3 mos. Avg.	3 mos. Change
-	1	0.5	◆ 130%
Page Views for rl-consult.com (what's this)			
Page Views per user:			
Today	1 wk. Avg.	3 mos. Avg.	3 mos. Change
-	14.0	11.7	◆ 103%

以上统计信息说明:www.rl-consult.com SEO 策略还是比较正确的,因为在几个月的时间内确实得到了较大的提升。同

(下转 274 页)

text中作为全局变量。用户发送的消息在保存到数据库的同时保存到接收者队列,接收者直接从自己的队列中获取最新接收到的消息,并清除已读取消息。在客户端,信息接收采用的技术为XMLHTTP:通过JavaScript脚本,在客户端建立XMLHTTP对象,然后利用该对象定时向服务端发出请求并取得最新接收的数据,将取得的数据重新组织并显示在客户端浏览器上,完成聊天信息接收页面的局部更新。也就是说,这里实现了一种高效,无页面刷新的B/S通讯方式。

系统可以动态显示在线用户和离线用户,方便员工的管理以及系统用户之间的交流。实现方法是为整个Web应用创建一个Session事件监听器UserListener,该类同时实现HttpSessionListener接口和HttpSessionBindingListener接口。在Web应用的配置文件web.xml中添加该监听器:

```
<Listener><listener-class>com.fwzx.web.UserListener</listener-class></listener>
```

创建一个单实例模式的类OnlineUser,该类包含一个保存在线用户信息的HashMap。在监听器中实现以下方法:

```
public void sessionCreated(HttpSessionEvent event){
    //session创建时调用,在此向HashMap中添加一个在线用户}

public void sessionDestroyed(HttpSessionEvent event){
    //session超时调用,从HashMap中清除超时用户,并清除该用户的消息接收队列}

public void valueUnbound(HttpSessionBindingEvent event){
    //用户注销时调用,从HashMap中清除注销用户,并清除该用户的消息接收队列}
```

4 运行效果

电子报税服务中心管理系统实际运行后,有效改进了旧系统不足,图3是系统运行时图。

运行效果达到了预定目标,包括:

(1)采用了上海CA中心(上海市数字证书认证中心有限公司)的CA证书和系统自身完善的用户管理模块,使得系统的安全性大大增强。

(2)采用了Hibernate的缓存机制,使得该系统与旧系统(C/S架构)相比,响应时间更短,提高了系统的服务效率。

(3)采用了J2EE架构和工作流的思想,系统工作流程和模



图3

块间结构更加清晰,扩展更加灵活,性能更稳定,对客户端硬件要求也大大降低,降低了客户端软件更新维护的成本。

(4)系统设计了对外接口,便于将来与将要开发的知识管理系统间的集成,使系统具有良好的对外开放性。

5 结束语

本文介绍了电子报税服务中心管理系统的设计与实现,论述了该系统基于J2EE框架的三层体系结构,介绍了本系统具有的各个功能模块及其具体设计,讨论了系统实现中的关键技术和关键模块的实现。相比旧系统,本系统综合运用了J2EE框架下的各项技术,使得该系统功能更加完善,工作流程更加清晰,技术先进,安全性大大提高,响应更加快速,系统的可靠性得到很大加强,并且提供了可扩展的接口,方便了与其它系统间的集成,系统运行效果达到了预定的目标。下一步将开发一个税务知识管理系统提供对该系统的支持,进一步提高该系统的效率。

参考文献

- 1.Ed Roman.Mastering EJB II[M].Wiley Computer Publishing,2002-01
- 2.孙卫琴.精通Struts:基于MVC的Java Web设计与开发[M].电子工业出版社,2005-02
- 3.James Goodwill.Mastering Jakarta Struts[M].Wiley Computer Publishing,2002
- 4.Budi Kurniawan.Java for the Web with Servlets,Jsp,and EJB:A Developer's Guide to J2EE Solutions[M].New Riders Publishing,2002-04

(上接255页)

时也说明,同是一个商业网站,是否进行搜索引擎优化,确实对网站的推广具有深远的影响。

6 其它SEO优化手法

目前,SEO优化方法纷呈复杂,手法各异,各大引擎排名参数众多,不一而论。除了上述影响因素外,还有很多,如:网站导航结构、网页表格的使用、title标签的写法、导入导出链接、内部链接、框架结构的使用、Flash优化、网页中CSS、JAVASCRIPT的应用、Google PR值研究、sandbox效应、Alexa世界排名等等。可见,SEO衍生出的搜索引擎优化服务领域有

着广阔的前景。

参考文献

- 1.Kai-Hsiang Yang,Chi-Chien Pan,Tzao-Lin Lee.Approximate Search Engine Optimization for Directory Service[C].In:Proceedings of the International Parallel and Distributed Processing Symposium(IPDPS'03),2003
- 2.马晓玲,吴永和.对于搜索引擎优化(SEO)的研究[J].情报杂志,2005,(12)
- 3.李福良,张辉.企业网站SEO技术研究[J].合肥工业大学学报(自然科学版),2004,27(1)
- 4.张红宇.浅论SEO及其实现[J].中国科技信息,2005,(16)