

从隐形网络到企业网站的 搜索引擎优化

谢 莉

(西南民族大学计算机科学与技术学院,四川成都 610041)

【摘 要】 本文从隐形网络的概念和成因入手,提出了面向用户的隐形网络观点,并从企业网站建设者的角度出发,指出为避免企业网站成为隐形网站,应利用搜索引擎优化技术对企业网站进行优化。

【关键词】 隐形网络;企业网站;搜索引擎优化

1 隐形网络与搜索引擎优化

1.1 隐形网络及其成因

隐形网络 (Invisible Web), 亦被称为隐蔽网络、深网或看不见的网络。Chris Sherman 和 Gary Price 在其合著的《隐蔽网络: 揭开搜索引擎看不到的信息源》一书中, 把“隐形网络”定义为: 虽然通过互联网可以获取, 但常规搜索引擎由于受技术限制而不能获得, 或者经审慎考虑后不作索引的那些文本页、文件或其他通常认为是高质量、权威的信息。显然, Chris Sherman 和 Gary Price 对隐形网络的定义是从搜索引擎的角度进行的。然而, 笔者认为, 隐形网络不仅是相对于搜索引擎而言, 而且是相对于搜索用户而言。因此, 本文把隐形网络定义为: 互联网上存在的但没有被常规搜索引擎收集到的, 或者是被收集到但由于结果排名等原因, 不易被用户获得的网络信息资源。

从当前研究看, 隐形网络形成的原因众多, 例如, 需要账号口令的数据库难以进入; 搜索引擎 Spider 查找的网页链接深度存在局限; 动态生成的网页难以搜索; 目前大部分搜索引擎只能索引 HTML 网页文档。从本文的定义出发, 结合当前互联网的发展状况, 笔者认

为, 目前隐形网络产生的原因还有以下几个方面:

(1) 搜索结果排名靠后。网络信息资源数量巨大, 一般搜索引擎检索的结果都有成千上万条相关或不相关的信息。虽然搜索引擎能够搜索到网站, 但是由于网站在搜索结果中排名太靠后, 用户可能没有足够的时间和耐心链接相应的网站, 从而使网站成为相对于用户来说的隐形网络。

(2) 未链接或错误链接的内容。由于搜索引擎的 Spider 或 Crawler 都是遍历网页上的链接收集网页的, 没有被链接的内容就等于没有给 Spider 和 Crawler 入口, 搜索引擎根本不可能找到此类网页, 这就形成了链接死角。

(3) 网站故意作弊。有些网站为了提高在搜索引擎检索结果中的排名, 使用一些搜索引擎拒绝的优化技术, 而被搜索引擎列为黑名单不予搜索, 从而成为隐形网络。

1.2 避免企业网站进入隐形网络

企业建立网站的目的是为了做宣传、进行网络营销, 或从事电子商务等。它们并不想让网站成为隐形网络, 而是希望搜索引擎能很容易地搜索到它们的网站, 并且在搜索引擎中得到较靠前的排名。

然而, 由于前面提到的隐形网络的种种成因, 都可

能让企业网站掉进隐形网络的陷阱。对于企业来说,不可能一味等着搜索引擎技术的改进来提高本企业网站的可获得性,也不可能靠用户自己改进搜索策略来发现企业网站。因此,为了避免成为隐形网络,企业的网站建设者们必须从自身网站的建设出发,对企业网站进行一系列的优化,使用户很容易地通过搜索引擎找到企业网站。

1.3 搜索引擎优化的概念和任务

搜索引擎优化(Search Engine Optimization, SEO)是指优化网站结构、网页代码和内容,使网站对全文搜索引擎友好,从而使网站易于被搜索引擎的蜘蛛程序找到,以提高网站在搜索结果中的自然排名。搜索引擎优化的主要任务是通过网站的、标签、排版等各方面的优化,使 Google 等常规搜索引擎更容易发现网站的内容,并且让网站的各个网页在常规搜索引擎中获得较高的评分,从而获得较好的排名。总之,好的搜索引擎优化工作能有效避免企业网站成为隐形网站。

2 企业网站的搜索引擎优化策略

2.1 结构平面化策略

搜索引擎理论上可以搜索到网站的任何一个层次的页面。但在实际搜索过程中,搜索 Spider 对于通过首页需点击 3 次及更多次才能进入的页面(三级页面及更深入的页面)不再重视,从而减慢、停止搜索,因此它们被搜索到的可能性很小。实际上,人们通常会通过点击三级页面上的链接才能进入浏览,而这些网页往往就被深埋于网站不为人知,成为隐形网络。因此,企业网站最好避免网页层次超过三层,尽量使网站的结构平面化。另外,企业网站结构平面化还可以通过创建网站地图实现,其中列出所有页面的链接,然后将网站地图链接到首页上(这也符合规范)。这样,企业网站中所有的页面都可以通过次级页面而被搜索到。

2.2 文本化策略

前文分析隐形网络形成的原因之一是网络资源的非文本化。目前,尽管很多搜索引擎开始支持多媒体检索(例如 Google 支持 .pdf、.doc、.ppt 等文件格式和图像检索),但目前多媒体检索技术还不成熟。要避免网站的资源成为隐形网络,目前最好的方法是将非文本资源文本化。其中包括:(1)在非文本元素(如图片、动画等)的代码中运用 Alt 属性标签进行说明,标签中要

包含关键词;(2)避免纯图像网页,如一些企业网站的首页形象页面;(3)少用 Flash,搜索引擎不仅不能理解 Flash 中的文字内容,也不能跟踪其内嵌链接;(4)对一些非文本元素的链接,可以在 URL 中列出其文本描述。

2.3 静态化策略

针对动态化的网页难于检索的特点,网站优化时,网站中重要的网页应尽量使用静态化网页。另外,还可以在一个静态页面(如网站地图)建立一个链接指向该动态页面,或者修改这个动态页面的 URL 使其中不再包含“?、=、&、%、+、\$”等符号。企业网站中很难避免使用动态网页,静态化策略并不能使所有动态网页都被检索到,但它可最大限度地使重要的动态网页内容被检索到。

2.4 关键词策略

搜索引擎描述每个网页的重要程度,最重要的因素之一就是关键词,这包括关键词的选择和关键词的密度。选择的关键词要有相关性,能准确定位网站的内容;要有较高的专指度,包括一些专有名词;热门度适中(太热不容易排前,太冷没人搜索)。搜索引擎抓取的文本中都尽量带有关键词,包括域名、title 和 meta 标签、正文、链接文本、文件名、alt、header 标签。所谓关键词的密度是指关键字与一个页面中除掉 html 代码内容的百分比。据一些专业人士观察研究表明,在大多数搜索引擎中,关键词密度在 2%~8% 是一个较为适当的范围,有利于网站在搜索引擎中的排名。关键词策略能使搜索引擎准确描述、索引网页的内容,从而提高网站在搜索结果中的排名,避免网站淹没在过多的搜索结果中而成为隐形网络。

2.5 链接策略

搜索引擎搜集和评价网页都是利用网页的链接。网络 Spider 是通过相关的链接来抓取网页的。搜索引擎的网页评级软件(如 Google 的 PageRank)也利用链接分析法来评价网页的重要性,搜索引擎将链接分析得出的网页级别与合适的关键词结合在一起,为用户找到最重要、最有用的网页。

要想让搜索引擎更容易找到你的企业网站,尽可能多地扩大链接的广泛度,让其他跟你主题相关的网站与你的网站链接,特别是那些重要网站的链接。另外,链接的深度也值得重视。很多网站的外链接(网站被其他网站链接)只是链接到网站的首页,如果能链接

到更深一层的页面,这会让你的网站为搜索引擎蜘蛛提供一个“侧门”。“侧门”可以让蜘蛛更容易地收录整个网站的页面。如果网站提供与主题相关的导出链接,被搜索引擎认为有丰富的与主题相关的内容,也有利于排名。

2.6 元标签策略

在网页的<meta> 标签中,对搜索引擎来说最重要的是关键词(keywords)和网页描述(description)。然而,除了 keywords 和 description 两个元素外,还可以根据一定的元数据集(例如 dc)设置更多的元素来描述网页的内容。但需要注意的是,目前只有 altavista 和 infoseek 两个主要的搜索引擎支持<meta>,有的 Web 设计者构建的<meta> 非常完美,但却发现自己的网站排在根本没有使用<meta> 的网站之后。也就是说,即使上述搜索引擎考虑<meta> 的作用,但与其他因素相比并不是最重要的。但笔者依然主张在网页中放置<meta> 标签,如果你不是滥用的话,对于你的网站排名,百利而无一害,而且相信今后会有越来越多的搜索引擎将<meta> 标签作为排名依据之一,从而使你的网站率先占据有利地位。

2.7 域名策略

域名中最好含有关键字,并且采用连字符“-”将该关键字单独突显出来以方便搜索引擎识别。有专家认为,域名中含有关键字对于排名的作用微弱,但不可否认确有作用,因此能够兼顾的情况下尽量考虑采用关键字域名。

2.8 避免惩罚策略

严格地说,避免惩罚并不应该称为策略,而是企业应该具备的某种道德。因此,企业应认真研究各搜索引擎明文禁止的 SPAM(垃圾)性质的优化技术,并给从事优化工作的技术人员进行职业道德培训,使他们做到绝对不用非法手段进行搜索引擎优化。

3 结 语

企业要想使自己的网站获得更多客户的关注,就需提升网站在搜索引擎中的自然排名。在企业的网络营销中,通过搜索引擎优化提升排名是一种非常重要的手段。使用搜索引擎优化服务得到的访问流量是免费的,从而可以源源不断地获得新的客户而不需要为点击付出额外的费用,这是一种作为长期目标来进行

的搜索引擎营销方式,但要达到理想排名的时间比较长,一般为 2 到 12 个月;采用竞价排名的好处是马上可以得到流量,但要为每个点击付费,排名越高,价格越贵,长期投入费用会很大,所以只能作为一种短期目标。因此,有学者建议,企业根据网站的实际来选择搜索引擎营销方式,如果是新网站,可同时选择搜索引擎优化和竞价排名来做搜索引擎营销;如果网站已有一定的流量,可以只选择搜索引擎优化来提高网站的流量,以达到较高的投资回报率。▲

参考文献

- 1 中国互联网络信息中心. 中国互联网络发展状况统计报告. <http://www.cnnic.net.cn/uploadfiles/pdf/2006/7/19/103651.pdf>
- 2 潘颖. 英文网站的搜索引擎优化及其海外宣传策略. 计算机系统应用, 2004(10)
- 3 娄卓男,吴志强. 近几年国外隐蔽网络研究概述. 图书情报工作, 2004(1)
- 4 吴雯洁. 企业的网站推广策略. 全球科技经济 望, 2005(6)
- 5 蔡东宏,李法运等. 网络信息资源开发与利用——从看不见的网站获取信息. 情报理论与实践, 2003(6)
- 6 梁焕平. 隐蔽网络及其检索策略研究. 情报杂志, 2004(7)
- 7 陈炎龙,张巨通. 从站点角度谈搜索引擎优化策略. 科技情报开发与经济, 2006(2)
- 8 C Brian Smith. Getting to know the invisible web. School Library Journal. New York: Summer 2001:16-18
- 9 Michael Dahn. Spotlight on the invisible Web. Online. Medford. Jul/Aug 2000, 24: 57-61
- 10 Anonymous. Indexing the“invisible web”. Information World Review. Oxford: Dec 1999:1
- 11 Glenn McGuigan. Invisible business information: The selection of invisible Web sites in constructing subject pages for business Collection Building. Bradford: 2003 (22):68-75
- 12 专家教你怎样做好中小企业网站搜索引擎优化. <http://www.jiajubiz.com/bbs/dispbbs.asp?board-id=13&id=2029&page=1>