

Otimização de *web sites* para sistemas de busca: uma estratégia de marketing online

Daniel Zander

Orientador: Cristiano T. Galina

Sistemas de Informação – Universidade do Vale do Rio dos Sinos (Unisinos)
São Leopoldo – RS – Brasil

dzpoa@hotmail.com

Resumo. *Este artigo apresenta um resumo do trabalho de pesquisa referente à otimização de web sites (SEO, acrônimo do inglês Search Engine Optimization) para atingirem as melhores colocações nos resultados apresentados pelos sistemas de busca. Apresenta os conceitos, quais fatores devem ser considerados, como planejar uma otimização, e um estudo de caso, onde serão aplicados os conceitos vistos.*

1. Introdução

Quando em 1990, na Universidade McGill de Montreal, Alan Emtage, criou Archie, a primeira ferramenta de busca usada na internet, o processo de busca pela informação era difícil e demorado. A internet era uma coleção de servidores *FTP* (*File Transfer Protocol*) interconectados que disponibilizava espaço para o *download* e *upload* de arquivos. Para encontrar uma informação desejada era necessário por parte dos usuários a navegação através de cada arquivo ou a indicação da exata localização por outros usuários. Archie armazenava e indexava em um banco de dados todas as listagens de diretórios de arquivos disponibilizados em dadas redes de servidores *FTP* anônimos, facilitando a sua localização pelos seus usuários [LEDFORD 2008].

Em 1991, Mark McCahill, da universidade de Minnessota, criou Gopher, um programa que indexava o conteúdo de documentos *plain-text*, que futuramente se tornariam os arquivos dos primeiros *web sites* da internet pública, época em que os servidores *HTTP* se tornariam mais populares que os servidores *FTP*. Gopher teve o apoio dos programas Veronica (*Very Easy Rodent-Oriented Net-wide Index to Computadorized Archives*) e Jughead (*Jozys's Universal Gopher Hierarchy Excavation and Display*) para referenciar os índices criados, fazendo buscas dentro dos arquivos armazenados, e assim permitindo aos usuários realizar buscas diretamente no índice, através da indicação de palavras-chaves [LEDFORD 2008].

Desde então, as buscas evoluíram e em 1993, Matthew Gray, criou Wandex, o primeiro sistema de busca que indexava e também buscava no índice de *web pages* da *web*. Sua tecnologia se tornou a base de todos os buscadores atuais, cujo software usa aplicações para coletar e armazenar em bancos de dados informações sobre *web pages*. Um sistema de busca possui uma interface onde seus usuários digitam um termo, uma palavra-chave ou frase, na tentativa de encontrar uma informação específica. Um algoritmo do sistema examina as informações armazenadas e retorna *links* de *web pages*

em que o conteúdo parece combinar com o termo digitado. A precisão dos resultados varia de acordo com o algoritmo usado [LEDFORD 2008].

O processo de coleta das informações referentes às *web pages* é realizado por um agente conhecido como *crawler*, *spider*, ou *robô*, que literalmente analisa cada *URL* (*Universal Resource Locator*) na *web* e coleta de cada *web page* palavras-chaves e frases-chave. O armazenamento e a recuperação dos dados depende dos métodos adotados por cada sistema de busca, e o *ranking* dos resultados mostrados depende dos critérios utilizados. Diferentes sistemas de busca usam diferentes critérios de *rankings*. Por esta razão uma pesquisa por palavra ou frase em diferentes sistemas de busca gera diferentes resultados [LEDFORD 2008].

O SEO foi idealizado como uma estratégia de marketing online que procura customizar os *web sites* a fim de manipular os sistema de busca, objetivando o melhor *ranking* possível, podendo assim gerar resultados consideráveis como aumento de tráfego, redução de custos e visibilidade através das primeiras posições nos resultados [THUROW 2003].

1.1 Motivação

Um grande percentual dos internautas, representado por 64% empregam “buscar” como o método primário de se encontrar conteúdo, serviços, produtos ou informação na web. Em 2004, os 3 principais motores de busca contabilizavam mais de 5% de todas as visitas a sites e mesmo quem nunca usou a internet ou “Googleou”, segundo [UNDERGOOGLE 2008] apresentado como uma expressão de verbo popular ao ato de pesquisar no Google, já ouviu falar a respeito do Google ou Yahoo!. No mesmo ano, 76% de todos internautas fizeram ao menos uma busca, totalizando 114 milhões de visitantes aos motores de busca [MORAN 2005].

Segundo [THUROW 2003] os consumidores tendem a confiar 5 vezes mais em sites que tenham encontrado em motores de busca e não em anúncios em banners, e que 28% desses consumidores digitam o nome do produto ou serviço no próprio motor de busca e não no site. Em 2006, um Instituto divulgou que cerca de 34% dos internautas consideravam que as principais marcas do mercado com relação a uma palavra-chave eram as que estavam apontadas nas primeiras posições pelo Google [VAZ 2008].

1.2 Objetivos

O objetivo deste trabalho é analisar as técnicas mais utilizadas em SEO, com o propósito de descrever uma estratégia, onde um site obtenha os melhores posicionamentos nos resultados das buscas do Google, Yahoo! e MSN Live Search.

Os objetivos específicos definidos são analisar o conceito de SEO em referências bibliográficas, identificar as técnicas mais importantes em SEO, planejar uma estratégia de otimização, implementar em um site já existente a estratégia planejada, e avaliar os resultados da implementação nos mecanismos de busca Google, Yahoo! e MSN Live Search.

2. A arquitetura dos *web sites*

A arquitetura dos *web sites* é o terreno no qual baseiam-se todos os esforços dos sistemas de buscas, como ilustrado na figura 1. Criar um *web site* com uma arquitetura otimizada passa a ser um dos principais fatores que contribuem para se obter e manter os melhores *rankings*. Uma arquitetura solidificada facilita os sistemas de busca a percorrer e entender o seu conteúdo [SIROVICH 2007].

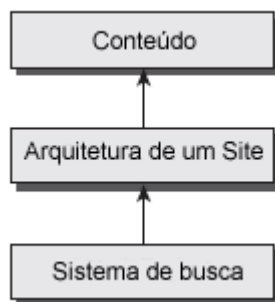


Figura 1. A arquitetura dos *web sites* como base dos esforços dos sistemas de buscas

Os efeitos de um bom conteúdo podem ser reduzidos substancialmente se algum aspecto ou tecnologias, como AJAX (*Asynchronous Javascript And XML*) ou *Javascript*, impedem ou confundem o acesso de um sistema de buscas às *web pages*. Os *crawlers* geralmente não seguem as instruções das *web pages*, não as executam, apenas as lêem. Sendo assim, as instruções devem, quando possível, ser executadas no lado do servidor e não no lado do cliente [SIROVICH 2007].

3. Fatores de *rankings*

Ao avaliar e comparar as *web pages*, todos os sistemas de busca trabalham a partir de diferentes algoritmos, e a importância dada a cada elemento interno e externo, que possa indicar o conteúdo existente em cada *web page* pode variar. A performance de um *web site* nos resultados mostrados é determinada pela maneira com que os elementos das suas *web pages* satisfazem os critérios de *rankings*. Estima-se que existam centenas de critérios diferentes em importância, onde os pesos e as medidas mudam constantemente. Os sistemas de busca não revelam seus próprios funcionamentos internos, porém os fatores fundamentais são conhecidos, compreendidos e garantem que um *web site* será devidamente indexado [SIROVICH 2007].

3.1. Fatores relacionados ao uso de palavras-chaves

Palavras-chaves são usadas para catalogar, indexar e encontrar um *web site*, já que potenciais visitantes as digitam nos sistemas de busca afim de encontrar *web sites* relacionados a um assunto específico. O uso de palavras-chaves populares e eficientes ajudam a assegurar que um *web site* será visível nos resultados dos sistemas de busca. Ao usar as palavras-chaves corretas, a base de visitantes será composta pelos que procuram exatamente o que o *web site* tem a oferecer [KENT 2006].

O uso da heurística torna-se importante em SEO na medida em que permite variações na maneira com que os visitantes buscam uma palavra-chave ou frase-chave.

Heurística é um termo que representa o reconhecimento de um padrão e a capacidade de resolução ou conclusão de um problema de forma rápida e eficiente através da consulta pelo que já se conhece do padrão, mesmo não sendo um conhecimento preciso. As palavras-chaves se aplicam às heurísticas pois provém o padrão pelo qual as buscas são feitas. Um sistema de busca opera de maneira heurística na hora de fazer as melhores escolhas e decisões no processo de *ranking* das *web pages*, concedendo pesos a cada palavra-chave encontrada de acordo com o seu conhecimento e experiência em relação ao assunto [LEDFORD 2008].

A escolha das palavras-chaves corretas pode ser feita através de uma análise ou checagem, de quais palavras ou termos as pessoas usam com maior frequência ao buscar na *web*. Esta análise pode começar com um *brainstorming*¹, identificando palavras óbvias e similares. Examinar *web sites* competidores e a origem do público visitante, através dos logs de acesso ao *web site* também ajudam a descobrir termos não pensados, como ilustrado na figura 2. Ferramentas especializadas, como a *Wordtracker*, a *Keyword Discovery*, ou a *The Yahoo! Search Marketing Research Tool* também ajudam no processo [KENT 2006].

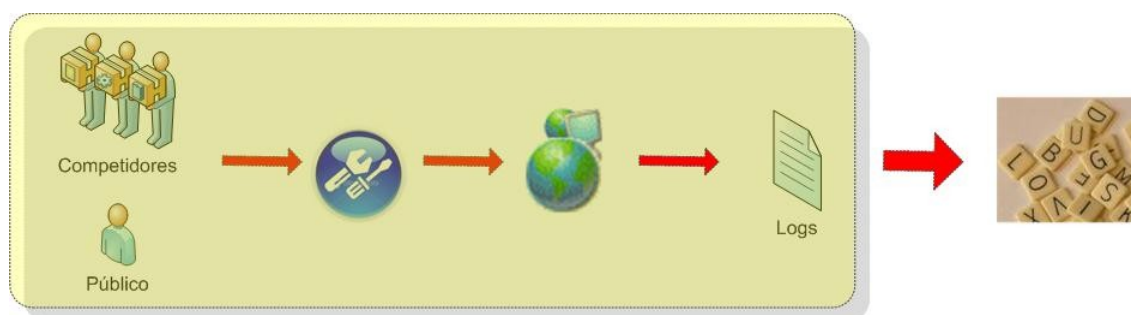


Figura 2. Descobrimo palavras-chaves

Um conceito importante, que deve ser levado em consideração, é sobre a porcentagem da quantidade de vezes que uma palavra-chave aparece na *web page* em relação ao total de palavras. Se muito alta, os sistemas de busca podem identificar que a *web page* foi feita apenas para chamar a atenção, e se muito baixa que a *web page* não é relevante ao termo buscado [KENT 2006].

Outro conceito importante é chamado *search tail*, e define que uma palavra-chave específica pode ser mais buscada que qualquer outra frase que a contenha, mas o total de todas as buscas pelas frases e não pela palavra-chave específica será sempre maior. Sendo assim, frases mais precisas e menos comuns podem ser mais efetivas em relação a palavras-chaves específicas e mais populares [KENT 2006].

3.2. Fatores relacionados aos atributos da *web page*

A forma com que os sistemas de busca interpretam dados específicos sobre as *web pages* também devem ser levados em consideração. Estudos mostram que a principal razão das pessoas visitarem *web sites* é a busca por informação, conteúdo. Para um sistema de busca, conteúdo significa texto, local onde ele identifica as palavras-chaves. Por isso, quanto mais texto um *web site* tiver, maiores serão as chances de suas

¹ **Brainstorming** – uma técnica onde um grupo de até dez pessoas se reúne e utiliza suas diferenças de idéias e pensamentos para assim gerar novas idéias [WIKIPEDIA 2008].

palavras-chaves serem identificadas e conseqüentemente aparecerem nos melhores resultados dos *rankings* [LEDFORD 2008].

Adicionar conteúdo em um *web site* não é uma tarefa simples e consome tempo, mas existem atalhos que podem simplificar o processo. Um dos atalhos, rápido e fácil, seria criar uma biblioteca de *links* onde, para cada link citado, existisse um resumo composto por muitas palavras-chaves. Outros atalhos seriam a construção de áreas como Q&A (perguntas e respostas), FAQ (perguntas freqüentemente respondidas), blogs ou seções de comentários [KENT 2006].

Certas *web pages* devem ser priorizadas ao serem adicionados conteúdos, criando assim pontos de entrada e/ou pontos centrais, para cada tópico do *web site*. Em SEO, páginas individuais são tão importantes quanto o *web site* inteiro. A frequência com que são atualizadas e a “idade” de cada *web page* também são levadas em consideração pelos sistemas de busca. As mais antigas são consideradas confiáveis, especialmente se elas são citadas em outros sites e não possuem *links* quebrados, enquanto as mais novas são temporariamente mais relevantes, pois tendem a possuir conteúdos atualizados [SIROVICH 2007].

A *URL* é o endereço de uma *web page* na internet que os visitantes digitam no *browser* ou que aparece nos *links* para que eles possam clicar. Os usuários tendem a clicar em *URLs* que se parecem mais com o termo procurado, então elas devem ser tão descritivas quanto possível, sem serem longas ou difíceis de serem memorizadas, chamando a ação e dando uma boa idéia do que será encontrado na *web page* a ser visitada, contribuindo assim com o aumento da taxa de cliques (*CTR*, *click-through rate*), outro fator considerado pelos sistemas de busca [LEDFORD 2008].

Imagens e gráficos em um *web site* são essenciais, porém são ignorados pelos sistemas de busca, pois os *crawlers* não conseguem indexar imagens, apenas textos. Porém, a técnica de inserir textos alternativos nas *tags HTML* de imagens, através do atributo *alt*, ajuda na obtenção de bons *rankings*. Os textos alternativos devem conter frases pequenas e descritivas a respeito da imagem, com a inclusão, quando possível, de palavras-chaves a respeito do conteúdo da página [KENT 2006].

A tolerância aos padrões e a limpeza de código não são importantes para os sistemas de busca, porém colocar as devidas *tags HTML* nos lugares certos é uma das maneiras mais efetivas de assegurar que um *web site* aparecerá nas páginas de resultados. Hoje as *tags HTML* incluídas em uma *web page* é muito mais relevante aos *crawlers* do que aos usuários. A formatação da apresentação aos usuários pode ser feita através de CSS (*Cascading Style Sheets*), deixando as *tags HTML* livres para fazerem seus trabalhos junto as *crawlers* [SIROVICH 2007].

O título de uma *web page* é visível tanto na barra de título do navegador *web*, como no texto do *link* dos resultados apresentados pelos sistemas de busca, apresentado na figura 3. Por isso, o título é um dos mais importantes fatores de *rankings*, pois chama a ação dos usuários e faz com que aumente a taxa de cliques, além de ser um dos primeiros elementos vistos e analisados pelos *crawlers*. Os títulos não devem ser genéricos, como nome de empresa ou slogan, e cada *web page* deve ter um título próprio e relevante ao seu conteúdo, contendo algumas palavras-chaves sem repetições [SIROVICH 2007].

[Search engine optimization - Wikipedia, the free encyclopedia](#)

Search engine optimization (SEO) is the process of improving the volume and quality of traffic to a web site from search engines via "natural" ("organic" or ...

[en.wikipedia.org/wiki/Search_engine_optimization - 93k - Cached - Similar pages](#)

Figura 3. *Link* de uma *web page* apresentado nos resultados de um sistema de busca

3.3. Fatores relacionados aos *links*

Os *links* sempre foram um fator importante nos *rankings* e recebem um grande peso junto aos sistemas de busca. São a principal rota de tráfego da internet e o alicerce de comunidades, agrupando *web sites* relevantes em conteúdo através de tópicos relacionados. Os sistemas de busca analisam e atribuem importância à relevância e à quantidade de todos os links de uma *web page* [LEDFOORD 2008].

Links devem habitualmente estar relacionados ao conteúdo da *web page*, e devem apontar para *web sites* relevantes às palavras-chaves, ou terão pouco valor para os sistemas de busca. Uma *web page* na qual um *link* aponta para um *web site* onde o conteúdo está relacionado recebe um peso maior nos *rankings*, porém o inverso também é verdadeiro. Apontar para uma “má vizinhança” ou *web sites* irrelevantes, faz com que a *web page* receba um peso menor [KENT 2006].

Quanto mais links apontam para um *web site*, maior importância ele terá nos *rankings* dos sistemas de busca. Os *links* internos de um *web site* garantem que os *crawlers* e os usuários irão navegar entre as suas *web pages*, enquanto que seus *links* citados em outros *web sites* contam como votos a seu favor. O sistema de votos funciona da seguinte maneira: os votos que surgem de *web sites* mais votados, portanto mais populares, recebem um peso maior em relação aos votos de *web sites* com menor número de votos. O algoritmo que aplica estas regras, nomeado *Pagerank* e ilustrado na figura 4, é implementado pela maioria dos sistemas de busca. Criar *links* relevantes contribui naturalmente para um alto *Pagerank* [LEDFOORD 2008].

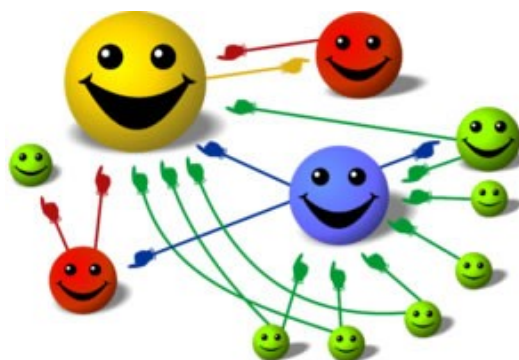


Figura 4. *Pagerank*

4. Planejamento em SEO

Com tantos fatores a serem considerados, um bom planejamento em SEO ajuda a visualizar onde os esforços precisam ser concentrados. Inicializado o planejamento, as atividades e as estratégias necessitam de um monitoramento e atualizações constantes.

Um plano em SEO deve ser considerado um documento dinâmico, sempre em mudança, apresentando a situação atual e o que precisa ser melhorado. A definição de um objetivo específico é necessária, de acordo com o negócio em questão, como por exemplo aumentar as vendas de algum produto ou apenas aumentar as visitas ao *web site*. Se um plano em SEO for estabelecido, o fluxo de tráfego ao *web site* tende a se tornar constante e balanceado, fazendo com que os *rankings* melhorem ao longo do tempo [LEDFORD 2008].

Ao invés de se olhar o site como um todo, cada *web page* precisa ser analisada. Algumas *web pages*, que por ventura geram maior número de tráfego de visitantes, devem ser priorizadas e o SEO pode ser planejado com base em cada uma delas, criando um roteiro de esforços de *marketing*. A avaliação deve ser feita por *web page*, documentando a situação de cada elemento, como *tags*, conteúdo, *links* e o mapa do site. A documentação irá conter colunas como nome do elemento, status, o que precisa ser melhorado, prazo para a melhoria e uma área para *follow-ups*, já que SEO é um processo contínuo [LEDFORD 2008].

5. Estudo de caso

O objetivo desta pesquisa será descrever a aplicação dos conceitos vistos, através de um estudo de caso único, no *web site* do Centro Comunitário Infantil Talitha Kum (<http://www.ccitalithakum.com.br/>), tendo em vista que o mesmo pretende aumentar o número de visitantes e a interatividade com os usuários.

O estudo de caso será aplicado no *web site* proposto, garantindo que ele seja suficiente para coletar os resultados esperados. Como ilustrado na figura 5, nas duas primeiras fase serão feitos um levantamento da situação atual do *web site*, a partir de uma entrevista junto aos proprietários, e a definição do objetivo que deverá ser alcançado ao se adotar SEO. Na fase seguinte será formulada uma estratégia de otimização, identificando as melhorias a serem realizadas, as páginas que deverão ser priorizadas, e quais as atividades de reformulação deverão ser postas em prática, criando assim um roteiro a ser adotado. A estratégia será fundamentada através de uma análise qualitativa sobre os fatores mais conhecidos em SEO, podendo assim classificá-los em ordem de relevância, de acordo com os elementos que serão empregados no *web site*. Com o planejamento aprovado e revisto, será iniciado a fase de implementação da reformulação do *web site*.

A avaliação do *web site* será feita constantemente para a coleta de resultados e monitoramento do tráfego, visitantes e classificações nos mecanismos de busca. Os dados coletados pelas avaliações deverão ser analisados no intuito de resolver o problema proposto (descrever um estratégia em SEO) e apresentar informações relevantes para uma conclusão.



Figura 5. Fluxo das Atividades a serem realizadas

Cronograma

- 1- Entrega do Artigo
- 2- Levantamento da Situação Atual
- 3- Definição do Objetivo
- 4- Formulação da Estratégia de Otimização
- 5- Implementação
- 6- Monitoramento
- 7- Elaboração da Conclusão
- 8- Escrita da Monografia
- 9- Entrega da Monografia

2008	2009						
DEZ	JAN	FEV	MAR	ABR	MAI	JUN	JUL
1							
2							
3	3						
	4						
	5	5					
6	6	6	6	6	6	6	
					7	7	
	8	8	8	8	8	8	8
							9

Tabela 1. Cronograma de atividades

Referências

- [LEDFOORD 2008] LEDFOORD, Jerri L. Search Engine Optimization Bible – Wiley Publishing, 2008.
- [SIROVICH 2007] SIROVICH, Jamie; DARIE, Cristian. Professional Search Engine with PHP: A Developer's Guide to SEO - Wiley Publishing, Inc., 2007.
- [THUROW 2003] THUROW, Shari. Search Engine Visibility - New Riders Publishing, 2003.
- [KENT 2006] KENT, Peter. Search Engine Optimization For Dummies, 2nd Edition - Wiley Publishing, Inc., 2006.
- [MORAN 2005] MORAN, Mike; HUNT, Bill. Search Engine Marketing, Inc.: Driving Search Traffic to Your Company's Web Site - Prentice Hall PTR, 2005.
- [VAZ 2008] VAZ, Conrado Afonso. Google Marketing: O Guia Definitivo de Marketing Digital, 2ª Edição – Novatec, 2008.
- [UNDERGOOGLE 2008] <http://www.undergoogle.com/blog/2005/11/vocabulrio-de-googlers.html>, Visualizado em 19/08/2008
- [WIKIPEDIA 2008] <http://pt.wikipedia.org/wiki/Brainstorming>, Visualizado em 03/12/2008