Jean Miguel Ferreira Dias

Otimização para motores de busca em um blog: Estudo de caso

Franca

2014

Jean Miguel Ferreira Dias

Otimização para motores de busca em um blog: Estudo de caso

Trabalho de conclusão para o curso de Sistemas de Informação no Centro Universitário de Franca – Uni-FACEF

Orientador: Leandro Borges

Franca

2014

Agradecimentos

Resumo

Abstract

Lista de Ilustrações

Lista de Tabelas

Lista de Abreviaturas e Siglas

SEO Search Engine Optimization

SEM Search Engine Marketing

HTML Hyper-Text Market Language

URL Uniform Resource Language

FTP File Transfer Protocol

Sumário

Introdução

Otimização para Mecanismos de Busca

SEO é uma sigla que significa Search Engine Optimization na tradução Otimização para Motores de Busca, que se refere a uma técnica responsável por fazer um website alcançar melhor visibilidade nas buscas, em ferramentas como o Google, Bing, Yahoo, entre outros.

Jerkovic (2009) define SEO como um agregado de todo o trabalho necessário para produzir um alto volume de referências bem-sucedidas oriundas de mecanismos de busca, diretórios web e outros websites, com o objetivo final de popularizar o website.

As técnicas usadas para melhorar o posicionamento de um website dentro dos grandes buscadores são várias, desde análises internas e externas do website, trabalho com mídias sociais, criação de links em sites de terceiros, otimizar conteúdo para palavras-chaves, criação de conteúdo e outras tarefas.

Mas antes de aprofundar em no tema do SEO precisamos entender alguns conceitos que servem como base para conseguir melhor consciência para implantar as técnicas de otimização em seu website. Com isso, para entender qual o motivo de otimizar um website, e quais fatores devem ser levados em consideração, é necessário entender os conceitos básicos do SEO.

Motor de Busca

Os motores de busca surgiram logo após a criação da Internet como conhecemos hoje. Segundo Ledford (2008) a primeira ferramenta de busca foi lançado em 1990, chamado Archie. A necessidade de se criar ferramentas que possibilitassem efetuar buscas dentro da Internet apareceu por causa da complexidade que era encontrar arquivos ou textos dentro da web.

A forma em que era encontrado um texto ou arquivos no começo da Internet não era nada parecido com o que nós temos hoje. Para você ter acesso a algo, a Internet baseava-se em um conjunto de pastas organizadas e para se ter acesso era necessário ter acesso por um protocolo de rede chamado de FTP (File Transfer Protocol, em português Protocolo de Transferência de Arquivos), onde permitia que aos usuários obter o acesso a estas pastas e procurar os arquivos que lhe sejam uteis.

Encontrar algo desta forma era bastante trabalhoso, pois consistia do usuário ter que navegar de pasta em pasta para achar a arquivo que o mesmo procurava. É lógico que existiam atalhos que levava o usuário até o seu destino rapidamente, mas era necessário saber onde encontrar estes atalhos, geralmente fornecidos quando você já sabia qual o arquivo exato você está procurando.

Com isso o processo para encontrar um arquivo na Internet para Ledford (2008) na prática, era um exercício difícil e demorado de paciência. Contudo em 1990 o estudante da universidade de McGill, em Montreal – Canadá, Alan Emtange criou a primeira ferramenta, que se tem registro, para busca na Internet chamada Archie, que era um índice com os arquivos da Internet.

A forma em que este programa denominado como Archie trabalhava está muito distante de como funciona os mecanismos de busca hoje em dia. Pois seu trabalho consistia em baixar uma lista de todos os arquivos dentro daquela pasta na Internet, e está lista era baixada para dentro de um banco de dados de websites, podendo assim ser consultado disponibilizando a informação de uma forma mais rápida.

O Archie servia apenas para listar arquivos de um FTP, não para procurar por arquivos relacionados a termo de busca, como acontece com os buscadores atualmente. Em 1991 outro estudante Mark McCahill da Universidade de Minnesota, percebeu que se existia a possibilidade de indexar arquivos, que era o que o Archie fazia, também era possível fazer busca de textos puro em referência de arquivos.

A partir disso como não existia nenhum aplicativo que realizava essa tarefa, ele criou o Gopher, um programa que permite indexar arquivos em texto puro, na figura 1 é possível ver como era o índice do programa, que segundo Ledford (2008) veio a se tornar os primeiros websites da Internet pública. Com isso a forma com que realmente funciona as buscas atualmente começou a amadurecer.

Mas para o Gopher funcionar com da maneira esperada foi necessária a criação de duas novas aplicações a Veronica (Very Easy Easy Rodent-Oriented Net-wide Index to Computerized, na tradução Índice Bastante Fácil para Arquivos Computadorizados em Rede) e Jughead(Jonzy’s Universal Gopher Hierarchy Excavationand Display, em tradução livre Escavação e Exibição Universal da Hierarquia do Gopher por Jonzy) que serviam para procurar os arquivos indexados nos índices do Gopher, denominado de Gopher Index System (na tradução Sistema de Índices do Gopher).

Figura 1: Índice do Gopher



Fonte: Wikipedia, 2008.

Mas só em 1993 desenvolvido por Matthew Gray aplicação denominada de Wandex, foi o primeiro programa a fazer tanto a indexação como também a busca nos índices das páginas da web. Segundo Ledford (2008) essa tecnologia foi o primeiro programa a vascular a web, e posteriormente se tornou a base para todos os crawlers de busca. No período de 1993 até o início do século 2000, foram desenvolvidos basicamente todos os motores de busca mais conhecidos do mercado, como o Yahoo em 1994, Google em 1997 e o MSN Search em 1998 que veio a se chamar Bing em 2008.

Estes três mecanismos de busca, Google, Yahoo e Bing são os mais populares do mercado na parte ocidental do globo, sendo o Google criado por Larry Page e Sergey Brin na Universidade de Stanford, em pesquisa feita pela comScore <http://goo.gl/IkS0uZ> em Março de 2014 nos Estados Unido, ele ocupa 67,5% das preferências de busca, seguido pelo Bing com 18,6%, Yahoo com 10,1% e outros com 3,8%.

Dentro deste mercado de motores de busca Jerkovic (2009) separa os mecanismos de buscas em alguns grupos, que são: motores de busca de primários, mecanismos de busca secundários, mecanismos de busca regionais, motores de busca verticais, mecanismos baseados em web spiders, motores de busca híbridos e baseados em metabusca.

Funcionamento de um Motor de Busca

No ponto de vista de um usuário final o motor de busca é apenas um simples website, que quando você precisa de procurar por algum termo na web você simplesmente digite essa palavra ou expressão em um campo, e a seguir pressiona a tecla “Enter” ou clica em pesquisar, onde que a partir destes pequenos passos vão surgir milhares de respostas para aquele termo digitado.

Mas essa definição fica por parte de um usuário final que não compreende realmente o que está acontecendo no back end do motor de busca. Vamos dividir o motor de busca em duas partes o back end e o front end, onde que no back end do motor contém a parte responsável por tudo funcionar e as buscas retornarem resultados, e o front end do motor apenas fica responsável por exibir para o usuário final estes resultados.

O back end de um motor de busca é apenas um software que segundo Ledford (2008) utiliza-se de aplicativos para coletar informações sobre páginas web. Estes aplicativos são chamados de crawlers, spiders ou robôs web. Ledford (2008) classifica estas informações coletadas como: palavras-chaves, expressões que represente o conteúdo da página web como um todo, a URL da página, e a estrutura dos links que saem e entram na página.

Com isso o motor de busca armazena estas informações a respeito de uma página web em um banco de dados. No front end que é a página web onde o usuário final digita seu termo para busca, o motor executa as seguintes tarefas: o motor irá pegar este termo digitado, ele fará uma consulta no banco de dados procurando se existe alguma página da web que contém aquele termo e após isso retornará todos os resultados encontrados exibindo-os ao usuário que efetuou a busca.

Tipos de Mecanismos

Neste capitulo será apresentado os vários tipos de mecanismos de busca presentes no mercado, onde que é possível classifica-los de acordo com sua abrangência e conteúdo retornado em suas buscas, referentes a localidade e itens retornados em seus resultados.

Mecanismos Primários

Para Jerkovic (2009) Bing, Google e Yahoo são mecanismos de busca primários, pois os três motores detêm maior fatia do mercado de pesquisas do mundo. Nesses tipos de mecanismo funcionam de forma que eles vasculham todas as páginas da Internet criando enormes banco de dados de índices, utilizados nas buscas.

Estes mecanismos de buscas primários são capazes de indexar a maioria ou até todos os websites da Internet, onde que Ledford (2008) fala que a diferença nos resultados deles fica apenas por conta de como o algoritmo do motor funciona, isto faz com que o mesmo termo procurado em um motor resulte em diferentes resultados em outro.

Mecanismos Secundários

Os mecanismos de busca secundários não têm o mesmo volume de acessos de um mecanismo principal pois geralmente são voltados para públicos mais específicos, isso faz com que o número de usuários que utilizam um mecanismo de busca primário seja maior que os motores de busca secundários.

No entanto mesmo Jerkovic (2009) apresentar que estes mecanismos são menos renomados e não tão populares quanto um mecanismo de busca primário, Ledford (2008) defende falando que para buscas regionais ou mais focadas eles podem ser bastantes uteis. Motores de buscas secundário podem ser citados como: Lycos, Ask, LookSmart, entre outros.

Mecanismos Regionais

Mesmo os mecanismos de buscas primários seja os quem geram maior trafego na Internet, existem os mecanismos regionais que são motores mais específicos para buscas em determinadas regiões, como o Baidu na China, que quando você utiliza-o seus resultados de busca são mais focados em páginas da Internet daquela região.

Mecanismos Vericais

Jerkovic (2009) chama-os de mecanismos de busca especializados, pelo fato de que os motores de buscas verticais pesquisam resultados mais específicos de um único tema, por exemplo um mecanismo de busca que contém apenas resultados para um único tópico como esportes, arte, ciências, literatura entre outros.

Mecanismos baseados em web spiders

Para Jerkovic (2009) os mecanismos de busca baseados em web spiders são definos assim:

Mecanismos de busca baseados em web spiders usam programas, o web spider, para criar seus índices de bancos de dados. Além do Google, Yahoo! e Microsoft, essa classe de mecanismo de busca inclui:

* Cuil (http://www.cuil.com)
* Exalead (http://www.exalead.com)
* Gigablast (http://www.gigablast.com)
* Teoma/Ask (http://www.ask.com)
* Walhello (<http://www.walhello.com>)

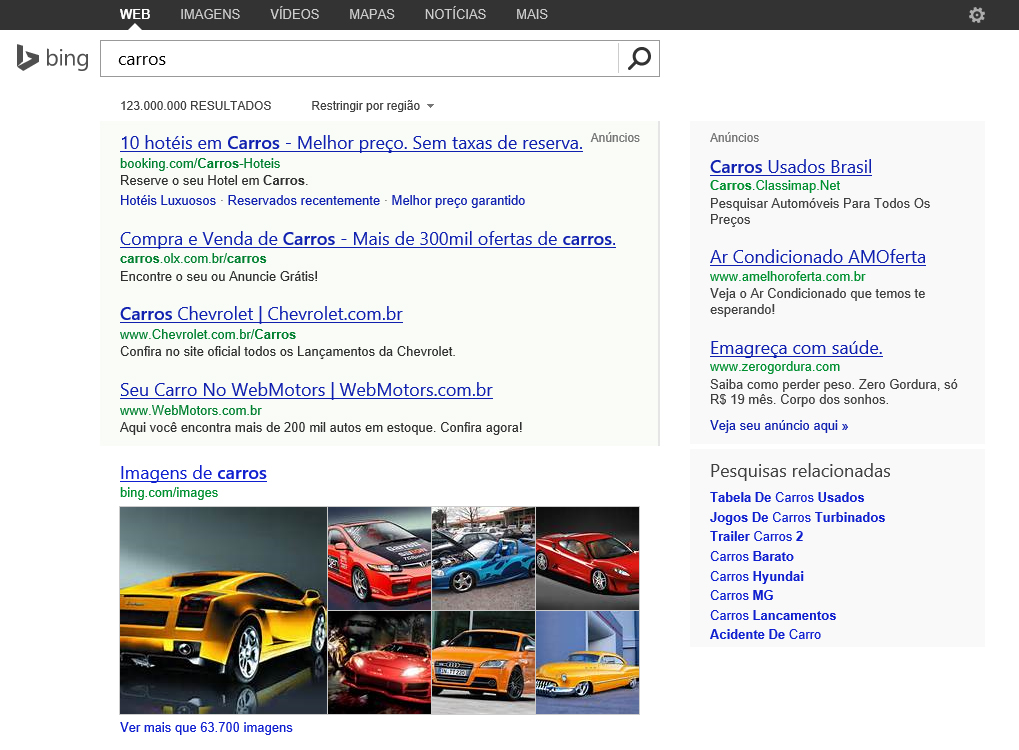
Mecanismos Hibridos

No mercado atual de buscas alguns mecanismos apresentam várias formas de efetuar pesquisa, por diversas categorias diferentes tal como ocorre nos motores de busca vertical. Estes mecanismos de busca são classificados como mecanismos de busca híbridos, que segundo Jerkovic (2009) é a unificação de uma variedade de resultados, onde é possível incluir resultados de spiders, diretórios web, notícias, produtos entre outros.

Motores de busca como Bing, Google e Yahoo! já adotaram este modelo de busca, que Jerkovic (2009) denomina de Busca Universal, onde que a ideias basicamente e informar múltiplos resultados para um termo, tornando assim a busca mais rica e completa. Na figura 1.2 é possível ver estes detalhes.

Na figura 2 é possível perceber como funciona um motor de busca hibrido, pelo fato da busca ser pela palavra “carro” é possível perceber que o mecanismo de busca retorna vários resultados diferentes, como os links de publicidade, uma área onde se encontra imagens sobre o termo procurado.

Figura 2: Resultados Variados do Bing



Fonte: o autor.

Mecanismos Metabusca

Jerkovic (2009) cita que se o usuário buscar pelo mesmo termo em diferentes mecanismos de buscas a base de spiders, eles lhe retornaram resultados diferentes, que ao contrário dos motores baseados em metabusca é organizar e juntar resultados de vários motores, fazendo com que assim aumente a qualidade nos resultados retornados.

Robôs

Os aplicativos responsáveis por alimentar o banco de dados de um motor de busca são conhecidos como Spiders, Crawlers ou Robôs Web, que tem como tarefa básica percorrer por todas as páginas web procurando por pelas informações que são armazenadas no banco de dados do motor de busca, Ledford (p7, 2008) explica o que são estes três programas da seguinte forma:

Essas pequenas criaturas são programas que literalmente vasculham a web, catalogando dados para que estes possam ser buscas pelos usuários. No sentido mais básico, todos os três programas – crawlers, spiders e robôs – são essencialmente a mesma coisa. Todos eles coletam informações sobre todas as URLs da web.

Com isso todas estas informações coletadas por estes três programas são catalogadas de acordo com sua URL (Uniform Resource Locator, na tradução Localizador Padrão de Recursos) para que assim possa ser encontrada após uma busca executada pelo usuário.

Jerkovic (2009) diz que o termo spider, robô e crawler representam a mesma coisa, no qual são aplicativos automatizados que tem como proposito de navegar pela Internet no objetivo de encontrar e fornecer ao seu mecanismo de busca o maior número de websites. Portanto nas próximas vezes que o texto se referir a spiders estará fazendo referência aos três de uma vez só.

Os Spiders atuam de forma aleatória dentro da web, eles navegam pelos websites rastreando não só os links mas também qualquer outro tipo de documento, como arquivos de imagens e vídeos. O acesso dos spiders aos websites podem ser várias vezes ou raramente, que segundo Jerkovic (2009) acontece a partir do tamanho do conteúdo do website.

Mas Jerkovic (2009) afirma também que os spiders não irão visitar seu website uma única vez, pois a partir do momento que um spider rastrear a página de seu website ele irá sempre voltar para detectar alterações, adições e exclusões dentro do conteúdo ou da página.

Os spiders de acordo com Ledford (2008) trabalham da seguinte forma: quando eles são lançados na web, eles recebem uma lista de websites que são o ponto de partida para os spiders, e a partir deste momento ele saem rastreando as páginas e encontrando novos links para percorrer.

Estes links que os spiders encontram podem ao mesmo tempo leva-los a outras páginas do mesmo website, como também podem leva-lo completamente para fora deste website. Com isso ele segue este processo até encontrar uma página do website que não leve para lugar nenhum, a partir deste ponto ele recomeça o processo até que todos os links sejam mapeados, segundo Ledford (2008).

Tipos de Busca

Dentro do cenário de SEO, muitos profissionais se confundem ele com o SEM (Marketing nos mecanismos de Busca) que são os anúncios de “pay-per-click” (Pague pelo Clique, PPC), por exemplo o Google AdWords. Embora ambos tenham alguns elementos em comum, como o foco em palavras-chaves, as implementações de SEM são bem mais simples, e alcançam resultados praticamente de forma imediata, pelo fato de serem formas de campanhas pagas, o SEO para Jerkovic (2009) é a forma mais próxima de uma propaganda gratuita que se pode chegar, pois estas técnicas lhe permitem aumentar sua visibilidade de forma natural, não necessitando pagar para aparecer como a primeira opção.

Mas entretanto não se pode dizer que para fazer SEO em um website não precisará desembolsar algum valor inicial, pelo fato do website estar posicionado nas buscas orgânicas, onde ao contrário do PPC, não necessita pagar para estar naquela posição. Fazer SEO exige que se faça investimentos, e seus resultados serão alcançados a longo prazo.

Jerkovic (2009) apresenta a seguinte tabela apontando as vantagens e desvantagens entre SEO ou PCC:

Tabela 1: Vantagens e Desvantagens SEO x PPC

|  |  |
| --- | --- |
| SEO | |
| Vantagens | Desvantagens |
| Benefícios a longo prazo | Resultados iniciais demoram |
| Aumenta a credibilidade pelo visitante | Para alguns casos não há garantias |
| PPC | |
| Vantagens | Desvantagens |
| Tráfego instantâneo | Gasta recursos rapidamente |
| Facilidade de implementação | As melhores posições são para quem paga mais |
| Facilidade de administração | Desconfiança |
| Menor investimento para manutenção | Visitantes param quando para o investimento |

Fonte: SEO Warrior

Com estes itens começa a perceber quais os grandes desafios dentro do SEO que podem ser elencados como a concorrência, falta de garantias, alterações constantes no ranking e fatores relacionados a tempo.

No mercado para cada setor você encontrará uma serie de competidores, onde que sua empresa, para obter um destaque aos olhos do consumidor necessita mostrar algum diferencial. No mundo virtual funciona da mesma forma existem vários concorrentes no mesmo segmento. Em pesquisa feita pela NetCraft em Abril de 2014, indica que existem 958 milhões de websites, 34 milhões a mais que no mês de março \url{http://goo.gl/zLvFD}.

De acordo com a ONU (Organização das Nações Unidas), em junho de 2013 a população mundial era de 7,2 bilhões, com este dado é possível afirmar que existe aproximadamente um website para cada sete pessoas \url{http://goo.gl/dcCDn6}.

Mas dentro de SEO segundo Jerkovic (2009) ninguém pode verdadeiramente garantir o primeiro lugar no Google, Yahoo! ou Bing. Pois para alcançar uma boa posição de *raking* dentro destes grandes buscadores, envolve vários aspectos que definem que seu website esteja em posições relevantes. Mas se você aplica todas as estratégias do SEO, tendo um conteúdo de qualidade, melhores posições e popularidade virão com o tempo.

Um outro ponto que afeta o posicionamento de seu website nos mecanismos de busca, fica por conta dos milhares domínios que estão competindo em busca de maior relevância, com isso ganha-se um aumento grande em sua concorrência. Por este motivo os buscadores vivem em constante melhoria para sempre estar retornando resultados relevantes em suas pesquisas.

Mas esse é um fator contempla dois lados, que é possível dizer que existe fatores positivos que auxiliam seu website a ganhar posições e melhorar dentro das buscas, e fatores negativos que servem apenas para atrapalhar seu website fazendo com que ele perda posições em páginas ou nem apareça nos resultados.

Um outro ponto que influencia no SEO é a questão do tempo, pois aplicar essas práticas e conceitos que melhora sua relevância, não é uma tarefa que lhe resultara em retorno a curto prazo. Ele lhe proporciona resultados a longo prazo e de forma natural. Isto faz com que em algum momento seu website comece a receber visitantes a partir de outros buscadores, além de outros websites que estarão vinculados ao seu.

Busca Orgânica e Busca Paga

Processos para Otimização de um Website

Neste capitulo serão abordadas as técnicas que devem ser utilizadas para que um website tenha sucesso ao realizar o trabalho de SEO, que suprem técnicas aplicadas na estrutura das páginas, conteúdo, relacionamento com parceiros e ferramentas que são uteis para poder acompanhar a evolução do website dentro dos resultados do Google.

Corpo da Página

Para aumentar a relevância de um website existe alguns processos que se deve analisar para realizar o trabalho de otimização e conseguir ter sucesso. Dentre estes processos estão os itens do corpo da página, conhecido também com itens *OnPage*. Estes itens são otimizações realizadas dentro da estrutura da própria concepção da página, no caso as tags de HTML.

Dentro da lista das diversas tags de HTML que contempla o corpo de website, existem algumas que tem influência direta para o SEO, na lista abaixo é apresentado as tags de HTML que apresentam maior relevância para o motor de busca.

* Title Tag;
* Tag A;
* Header Tag (H1, H2...H6);
* Tag P, Article, Section, Nav, Header e Footer;
* Tag Strong.

Links Internos e Externos

Conteúdo

Ferramentas de Analise

Visando sempre aprimorar o posicionamento de seu website, fazendo com que ele sempre esteja legível para o Google e além disso possa acompanhar uma análise detalhada para que seja possível avaliar os pontos fortes e pontos fracos de seu website, a própria Google desenvolveu duas ferramentas que são de extrema importância para o administrador do website possa ter informações e estatísticas, o que permitirá a tomada de decisão para gerar mais visitantes.

Essas ferramentas são conhecidas com Google Webmaster Tools e Google Analytics. A primeira ferramenta segundo Jerkovik (2009) permite aos websmaster um meio fácil de rastrear o tráfego de suas páginas, também analisar o arquivo robots.txt, adicionar seus arquivos sitemaps e assim por diante. E a segunda ferramenta o Google Analytics segundo Jerkovic (2009) é a ferramenta de marketing preferida para a visualização de estatísticas web gerais e o mais importante, o acompanhamento de conversões.

Para utilizar ambas as ferramentas, é necessário possuir uma conta da Google para que assim possa ter acesso ao Google Webmaster Tools e o Google Analytics. Ambas as ferramentas para que possam funcionar de acordo dentro do website e começem a capturar as informações do website para assim exibir para analise, elas precisam primeiro reconhecer o website, tarefa no qual pode ser feita tanto subindo um arquivo com o nome gerado randomicamente pelo Google na raiz do website, como também a partir de uma inserção de código nas páginas do website.

Ferramenta para o administrador do website

O Google Webmaster Tools é uma ferramenta onde é possível administrar vários websites, com isso em sua tela principal aparece uma lista com todos os websites que você administra. No painel esquerdo da ferramenta a partir do momento em que se é selecionado um website, é apresentado um menu com acesso a todas as ferramentas para administração e analise de seu website. Na tabela 2 é apresentado uma lista com todas as ferramentas disponíveis no Google Webmaster Tools.

Tabela 2: Ferramentas do Google Webmaster Tools

|  |  |
| --- | --- |
| Ferramenta | Descrição |
| Painel do Site | Mostra uma visão geral do website |
| Mensagens do site | Mostra todas as notificações ou alertas do website |
| Dados estruturados | Mostra uma visão geral se o Google identificou dados estruturados no website |
| Marcador de dados | Permite que você marque os dados estruturados no website |
| Melhorias de HTML | Informa setores do website que podem ser otimizados |
| Links para o site | Permite rebaixar os sitelinks do website |
| Consultas de pesquisa | Permite visualizar quais são as pesquisas que os usuários fazem e acessam o website |
| Links para seu site | Permite visualizar quais são os links de acesso para o website |
| Links internos | Permite visualizar quais os links internos do website |
| Ações manuais | Permite visualizar as ações manuais contra spam no website |
| Segmentação internacional | Permite visualizar para qual público alvo em nível internacional o website |
| Status do índice | Permite visualizar o status do índice das páginas do website dentro dos resultados do Google |
| Palavras-chave do conteúdo | Permite visualizar quais são as palavras-chaves que aparecem no conteúdo do website |
| Remover URLs | Permite remover URLs do website de dentro dos índices do Google |
| Erros de rastreamento | Exibe todos os erros de rastreamento de páginas do website pelo Google |
| Estatísticas de rastreamento | Exibe um relatório de rastreamento do website pelo Googlebot |
| Buscar como o Google | Permite que possa ser procurada páginas no website com o próprio Googlebot |
| Testar robots.txt | Permite que possa ser testado o arquivo robots.txt do website |
| Sitemaps | Permite enviar e verificar quais as URLs do sitemap foram processadas pelo robô do Google |
| Parâmetros de URL | Permite adicionar parâmetros para melhor rastreamento das URLs do website |
| Problemas de Segurança | Informa se o website apresenta algum problema relacionado a segurança tal como vulnerabilidades |
| Outros recursos | Apresenta sub-ferramentas da própria Google para auxiliar nas melhorias e divulgação do website |

Fonte: o autor.

No menu Painel do site na ferramenta Google Webmaster Tools ela apresenta algumas informações mais importantes para análise do website como, informações de status de rastreamento do website, consulta de pesquisas, sitemap e algumas mensagens críticas. Todos estes dados de forma genérica mostrando as informações de uma forma simples sem todos os detalhes que são apresentados em seus respectivos menus.

Na seção Mensagens do site é apresentada todas as mensagens de erros, alertas e notificação sobre o website, onde que o próprio robô de busca do Google, o Googlebot, vai identificando em seu rastreio e armazenando e relatando neste painel, para que o webmaster fique informado sobre as situações do website e posso sempre estar corrigindo e melhorando.

A ferramenta de Dados estruturados segundo o Google (2014a) mostra ao websmaster informações estruturadas que o Google detectou em seu site, esta ferramenta permite que o Google analise o website e consiga encontrar as marcações de dados estruturados, o que faz o website ser melhor compreendido pelo motor de busca, e em pesquisas efetivadas pelo usuário o Google possa retornar os seus resultados com melhor precisão.

Na ferramenta Marcador dados é possível selecionar e marcar todas as seções e itens do seu website, para que o Google no momento em que for percorrer pelo website dentro de seu índice, ele automaticamente irá reconhecer as marcações realizadas com o uso da ferramenta, e apresentará o website de uma forma mais agradável exibindo mais itens e informação no resultado.

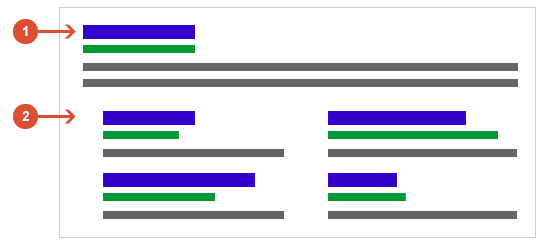
Melhorias de HTML é uma ferramenta que permite visualizar problemas encontrados pelo Google durante o rastreamento e a indexação do website. O Google recomenda que seja acessado este painel para identificar os problemas que se forem corrigidos auxilia a aumentar a classificação do website nas páginas de resultados. Segundo o Google (2014b) estas questões não evitam que o site seja rastreado ou indexado, mas cuidar delas permite que o website possa oferecer maior experiência ao usuário e também ajudar a direcionar tráfego ao website.

Entre os dados contidos nas melhorias de HTML o webmaster encontrará informações como: problemas com os títulos das páginas, por exemplo títulos repetidos nas páginas, problemas com a tag meta descrição, podendo ela estar vázia ou duplicada em outras páginas, e até páginas com conteúdo que não pode ser indexado como arquivos de formatos desconhecidos pelo robô de busca.

Na ferramenta Links para o site é permitido que o webmaster possa rebaixar os links para o site que não seja interessante para o usuário, no caso esses links segundo o Google (2014c) são chamados de sitelinks, que tem a finalidade de ajudar o usuário a navegar pelo website com maior facilidade.

Estes sitelinks são exatamente os links exibidos abaixo do link retornado pela busca, na figura 3 é exibido no item 1 o link principal do website e no item 2 os sitelinks. Os sitelinks exibidos pelo Google apenas são apresentados caso a estrutura do website permitir, caso eles não sejam pertinentes a consulta do usuário eles nem serão exibidos. Segundo o Google (2014c) ele identifica os sitelinks a partir do texto âncora e atributo alt dos links internos do website.

Figura 3: Webmaster sitelinks



Fonte: Google, 2014c

Consultas de pesquisa é a página dentro do Google Webmaster Tools em que é possível visualizar as consultas efetuadas no Google, em que as mesmas retornam páginas do website, nesta ferramenta é possível ter dados como a taxa cliques (CTR) no website para cada termo procurado, como também a posição média da página dentro dos resultados.

Na ferramenta Links para seu site é a parte do Google Webmaster Tools onde ele classifica todos os links de websites de terceiros que apontam para o website que está sendo gerenciado. Dentro da ferramenta de busca do Google, segundo o mesmo (2014d), ele indica que se for digitado na pesquisa o termo por exemplo "link:tccseo.com.br", a busca retornará páginas da web que aponta para o website.

Na área de Links internos, a ferramenta mostra os links do website que possuem entradas internas para páginas do próprio website, que no caso são sub-links para páginas internas.

Ações manuais é uma página do Google Webmaster Tools que tem como objetivo apresentar ações manuais realizadas pelo próprio Google afim de previnir que em seus resultados apareçam websites que utilizam de técnicas de spam. O Google (2014e) cita que mesmo que seus algoritmos sejam sempre atualizados para previnir isso, ele podem utilizar técnicas manuais para rebaixar ou até mesmo excluir a página dos resultados.

Embora o Google use algoritmos para avaliar e aprimorar regularmente a qualidade da pesquisa, também estamos dispostos a realizar ações manuais em sites que usam técnicas com spam, como rebaixar ou até remover esses sites de nossos resultados de pesquisa completamente.

O Google (2014e) informa que caso seu website ou uma página não apareça mais nos resultados, deve-se consultar a página de Ações manuais para ver os procedimentos necessários para fazer a página voltar, e que ao realizar tais procedimentos é possível solicitar ao Google uma análise diretamente do website ou da página em questão para que ele volte aos resultados.

Na página de Segmentação Internacional para o administrador do website o Google permite que caso o website tenha versões diferentes para cada idioma ou região, é possível tomar medidas para que o conteúdo em específico do website alcance o público-alvo para aquela página. Para fazer isso o Google (2014f) explica que basta adicionar uma tag de HTML no cabeçalho do website para que ele identifique qual a página recomendada para cada região. No trecho a seguir o Google (2014f) explica três formas para fazer essa marcação.

Para isso, é possível usar três mecanismos de implementação:

* **Marcação no nível da página**: Use a tag <link rel='alternate' hreflang='x' href='alternateURL'> na seção <head> das suas páginas para listar as versões alternativas de idioma de cada uma delas. Cada página deve fornecer uma tag hreflang que vincule a todas as suas outras variantes de idioma, assim como uma tag de referência que direcione para a página original. Para uma segmentação mais granular, use o atributo hreflang para indicar as combinações de idioma e país (por exemplo, en-ie, en-ca, en-us). Leia mais sobre a tag hreflang na seção "Diretrizes de conteúdo".
* **Sitemaps**: Use Sitemaps para enviar opções alternativas de idioma e região para suas páginas. Leia mais sobre o uso de Sitemaps para indicar as páginas com idiomas alternativos na seção "Diretrizes de conteúdo".
* **Cabeçalhos HTTP**: Se você publicar arquivos que não sejam HTML, como PDFs, use um cabeçalho HTTP para indicar uma versão em um idioma diferente de um URL.

A seção Status do índice dentro do Google Webmaster Tools, segundo o Google (2014g) ela fornece estatísticas sobre os URLs indexados pelo motor do Google para o website em questão.

Palavras-chave de conteúdo é uma área da ferramenta do Google que possibilita o administrador do website por visualizar as palavras-chaves que a própria Google identifica dentro do conteúdo do website, com esta informação é possível direcionar os esforços em SEO para algumas palavras-chaves que o Google identifica que sejam mais relevantes no próprio website.

A ferramenta Remover URLs dentro do Google Webmaster Tools é uma ferramenta que permite que o administrador remova determinadas URLs que estejam indexadas pelo Google mas que não sejam necessárias estarem lá. Com isso é possível enviar uma solicitação de exclusão de URLs do index do Google, a própria Google sugere que você especifique no arquivo robots.txt quais seção do website o administrador do mesmo deseja que não seja indexado.

Na página Erros de rastreamento o Google informa ao administrador do website quais páginas do website estão com algum erro quando o robô de busca tenta indexar a mesma. Os erros mais frequentes dentro desta seção são os erros 404 e 500 (PRECISO DE UMA REFERENCIA EXPLICANDO OS DOIS ERROS). Com estas informações disponibilizada nesta página é possível identificar e corrigir as páginas que não estão sendo indexadas é informar ao Google a correção do erro para que seu motor de busca possa tentar indexa-lo novamente.

A página de estatísticas de rastreamento fornece gráficos informando alguns dados sobre a atividade do Googlebot no processo de indexar as páginas do website. Nesta área é informado dados como atividade de páginas rastreadas por dia em um período de 90 dias, além quilobytes recebidos pelo Googlebot do website é o tempo de resposta em milissegundos do website.

Na ferramenta Buscar como o Google do Google Webmaster Tools permite que administrador do website possa fazer busca de páginas do website como se fosse o próprio Google executando a busca afim de indexar a página, segundo o Google (2014h) este recurso é bastante útil caso o website apresente uma página que tenha problemas de desempenho ao retornar nos resultados, pois ela mostra o conteúdo que o robô do Google retorna ao solicitar a página ao servidor.

O arquivo robots.txt é um arquivo que segundo o Google (2014i) impede um software rastreador da Web, como o Googlebot, de rastrear determinadas páginas de seu site, na página Testar robots.txt é possível que o administrador do website possa editar o e identificar possíveis erros dentro do arquivo robots.txt, afim de fazê-lo funcionar de forma correta, indicando corretamente as regras para o Googlebot.

Na página Sitemaps é possível enviar o arquivo de sitemap do website para o Google, o que possibilitará com que todas as páginas contidas dentro do mesmo sejam indexadas pelo Google. Nesta página também indica se o Google encontrou algum erro no arquivo de sitemap ou se ele ainda está pendente ou se já foi lido pelo robô do Google.

O Google (2014j) define que seu objetivo é indexar seu website da maneira mais eficiente possível, e que rastrear conteúdo duplicado ou idêntico é um uso ineficaz de seus recursos, com isso na página da ferramenta Google Webmaster Tools, Parâmentros de URL é possível adicionar algumas regras que possam ajudar o Google a identificar parâmetros dentro das URLs de seu website de forma mais inteligente, evitando rastreio de conteúdo duplicado.

Na página Problemas de segurança o Google informa para o administrador do website se o mesmo apresenta algum problema de segurança, como por exemplo malware que segundo Microsoft (2014) são quaisquers softwares indesejados, instalados sem a devida permissão. Estes problemas podem afetar o funcionamento do website podendo até prejudicar o visitante das páginas, por isso nessa seção o Google faz uma identificação e mostra os problemas que ele encontrou na segurança do website.

Na seção de outros recursos o Google disponibiliza algumas ferramentas que permitem auxiliar os processos de melhorias no website. São um total de sete ferramentas que são:

* Ferramenta de teste de dados estruturados tem como finalidade testar dados estruturados para verificar se o Google consegue analisar corretamente os marcadores de dados estruturados e exibi-los nos resultados quando o website aparecer em alguma pesquisa.
* Assistente de marcação de dados estruturados.
* Email Markup Tester ferramenta para validar os conteúdos de dados estruturados de um e-mail em HTML.
* Google Places permite que seja informado ao Google informações da empresa para que ele retorne estes dados dentro no momento da busca.
* Google Merchant Center é o local para disponibilizar dados de loja e dos produtos para que sejam exibidos no Google Shopping.
* PageSpeed Insights informar dados para o administrador do website poder tornar as páginas do website mais rápidas em todos os dispositivos.
* Pesquisa personalizada neste recurso o Google permite que dentro do seu próprio website tenha um caixa de pesquisa do Google que irá utilizar do motor da busca da Google para fazer pesquisas dentro do website.

Ferramenta para o profissional de marketing

Aplicação

Este trabalho tem como objetivo apresentar na prática os conceitos e técnicas para otimização de um website para os mecanismos de busca, onde que no estudo de caso apresentado o mecanismo em questão será o Google. Dentro deste ambiente será colocado em prática todas as técnicas já descritas para que o website em questão possa conseguir uma boa classificação dentro de uma busca em específico.

Como tema para o website para o ser utilizado como estudo de caso do projeto foi escolhido um blog com o tema sobre SEO, onde que este blog é referente ao tema deste trabalho. O objetivo da escolha de um blog para ser o website foi em cima do fato de que dentro de um website deste gênero é possível explorar todos os conceitos e técnicas para otimização de um website.

No desenvolvimento deste website que pode ser encontrado através do endereço de web *www.tccseo.com.br* foi utilizado tecnologias de desenvolvimento como o framework para aplicações web Ruby On Rails na versão 4.0.3 com sistema de gerenciamento de banco de dados o MySQL na versão 5.5, ambas as ferramentas têm ampla compatibilidade entre si possibilitando um desenvolvimento ágil desta aplicação.

A modelagem do banco de dados está disponibilizada no anexo 1, em que é possível observar que a base de dados constitui de apenas 4 tabelas que permitem que o website trabalhe de forma satisfatória, que são: Usuários, Categorias, Calendário e Artigos. A modelagem do banco de dados foi utilizada o modelo de entidade de relacionamento (MER), criado por Peter Chen em 1976.

O MER de acordo com Silberschatz (2006) foi desenvolvido para facilitar o projeto de banco de dados, permitindo especificações de um esquema organizacional que representa a lógica de um banco de dados. A partir deste modelo temos o DER (Diagrama de Entidade de Relacionamento) que permite demonstrar de forma gráfica a estrutura lógica do banco de dados.

O anexo 2 apresenta os requisitos analisados para o desenvolvimento do website, neste anexo é possível visualizar os requisitos tal como o ator responsável pela execução do mesmo. A partir destes requisitos foi possível desenvolver a UML (Unified Modeling Language, em tradução livre Linguagem de Modelagem Unificada) desta aplicação que pode ser encontrada no anexo 3.

A UML de acordo com Guedes é uma linguagem visual para modelar sistemas por meio do paradigma de Orientação a Objetos. A partir desta premissa o Ruby linguagem de desenvolvimento criada por Yukihiro Matsumoto em 1995 segue o mesmo paradigma de orientação a objetos o torna a tradução da UML para Ruby totalmente portável.

Em 2004 a partir da linguagem de desenvolvimento Ruby uma *startup* chamada 37signals criou um framework ágil para desenvolvimento web, denominado Ruby On Rails que segundo Hartl (2012) tornou-se uma das mais poderosas e populares frameworks para desenvolvimento de aplicações web dinâmicas.

O uso do MySQL como banco de dados foi escolhido pelo fato de ser um dos mais populares mercados de acordo com Kofler (2005) é fato que o número de usuários do MySQL vem crescendo a cada dia, e uma maioria esmagadora está bem satisfeita com ele, mas não pode-se afirmar que ele possui mesmos recursos tal como seus concorrentes comerciais, por exemplo SQL Server e o Oracle.

O Website desenvolvido para estudo de caso deste trabalho como já havia sido comentado é um blog que tem como objetivo informar sobre termos, técnicas e conteúdos abordados neste trabalho, mas de uma forma informal para fácil compreensão de um público variado.

Este website que nos próximos paragrafo também será utilizado o termo blog para referenciar ele, é um website da categoria do tipo blogue, no PRIBERAM (2013) define como página de internet com características de diário, atualizada regularmente, esta categoria de website permite que você utilize melhor o conceito abordado no subcapítulo sobre Conteúdo.

O fato da escolha do estudo de caso ter sido escolhido um website da categoria blog e não um website do tipo institucional, loja virtual, fórum ou websites de entretenimento, deve-se ao fato de qualquer um dos modelos que foram apresentados a implementação de SEO acaba pecando em algum momento.

No caso de um website do institucional é possível apresentar todas as técnicas, mas o fator conteúdo acaba ficando limitado fazendo com que seja criado um blog. O mesmo acontece para os demais tipos de websites, lojas virtuais, fóruns e websites de entretenimento, como já foi citado na subcapítulo sobre Conteúdo, este é um dos principais fatores que o Google avalia para aumentar a relevância de um website.

Cada aspecto para a construção desde website do tipo blog foi calculado para que todas as técnicas em conjunto tenham sucesso. Nos critérios de corpo da página as estruturas de tags foram feitas de acordo com a semântica mantendo o a hierarquia em que alguns websites não existem ou são colocados de forma aleatória não respeitando a hierarquia.

Para melhor identificação do conteúdo do blog pelos motores de busca, foram adicionados ao decorrer de todo o código-fonte as tags de *rich snippets* do website Schema.org, onde que foram utilizadas três seções de sua hierarquia de referência: o *Blog*, *Website* e *Events*.

Com a utilização destas três seções foi possível melhorar a identificação do conteúdo do website por parte dos mecanismos de buscas, como já foi citado no subcapitulo sobre Corpo da Página. Toda a estrutura de links do website foi trabalhada para que o visitante navegue pelo website com tranquilidade e quando clicar em um link de parceiro ele não saia complemente da página.

Os links foram trabalhados desde seu conteúdo chave, escolhendo sempre a melhor palavra-chave para poder utilizar levando em consideração o conteúdo da página e os termos que o visitante costuma a buscar para encontrar aquele conteúdo que será apresentado a partir do link. Também com relação aos links foram analisados os textos que iram fazer parte da propriedade de título para o mesmo, pois nesta propriedade é uma forma de informar algo a mais para aumentar a relevância da página destino do link.

No critério conteúdo pelo fato do tipo do website ser um blog gerar conteúdo deixa de ser um problema, pois a natureza de websites deste tipo é conter conteúdo. Todos os textos publicados correspondem diretamente ao conteúdo apresentado por este trabalho, mas escrito de forma informal para alcançar todos os públicos.

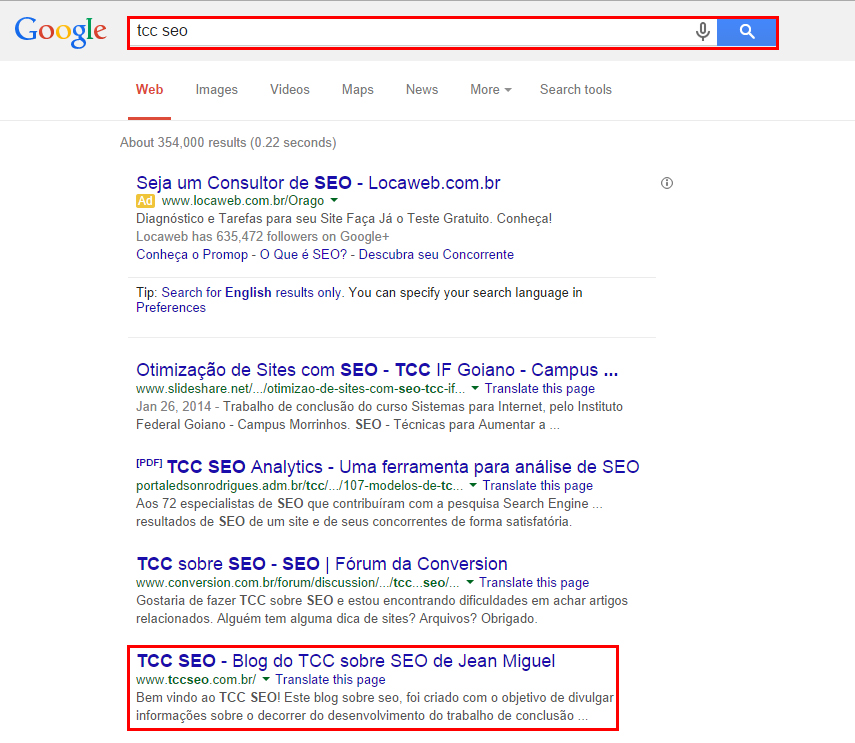
Análise de Evolução

A evolução do website foi acompanhada diária através da ferramenta Google Analytics e Google Webmaster Tools, onde que foi apresentado dados em relação a evolução do tráfego de visitantes, tanto novos, quanto visitantes que retornaram ao website.

Também com auxilio destas ferramentas foi possível acompanhar quais páginas já tinham sido indexadas aos resultados de busca, no caso os do Google, além de quais palavras-chaves elas estavam aparecendo e quais suas posições médias, tal como também taxa de clique (CTR) para cada termo procurado.

O Blog TCC SEO desde do período que foi publicado no dia 23 de Julho de 2014 até o atual momento, segundo a ferramenta Google Webmaster Tools sua posição média era o oitavo lugar na busca por "blog do tcc", no entanto para a busca "tcc seo" o website já se apresenta na quarta posição atrás apenas por outros trabalhos sobre SEO apresentados anteriormente. Como pode ser visto na figura 4, a posição atual do website nos resultados do Google.

Figura 4: Posição atual do Blog



Fonte: o autor.

Conclusão

Referências bibliográficas

JERKOVIC, John I. SEO Warrior. O'Reilly Media, 2009.

LEDFORD, Jerri L. SEO Search Engine Optimiztion Bible. 1.ed. Starlin Alta Con, 2008.

HARTL, Michael. Ruby On Rails Tutorial. Learn Web Development with Rails. 2.ed. Addison-Wesley, 2012.

GUEDES, Gilleanes T.A. UML 2. Guia Prático. Novatec, 2007.

SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN,A. Sistema de Banco de Dados. 5.ed. Mc-Graw-Hill, 2006. Tradução por Daniel Vieira. Rio de Janeiro Elsevier, 2006.

KOFLER, Michael. The Definitive Guide to MySQL 5. 3.rd. Apress, 2005.

GOOGLEa. Monitorar a marcação de dados estruturados. Disponível em: <https://support.google.com/webmasters/answer/2650907?hl=pt-BR>. Acesso em: 19/08/2014.

GOOGLEb. Melhoriasde HTML. Disponível em: <https://support.google.com/webmasters/answer/80407?hl=pt-BR80407>. Acesso em: 21/08/2014.

GOOGLEc. Sitelinks. Disponível em: <https://support.google.com/webmasters/answer/47334?hl=pt-BR>. Acesso em: 21/08/2014.

GOOGLEd. Links para seu site. Disponível em: <https://support.google.com/webmasters/answer/55281?hl=pt-BR55281>. Acesso em: 23/08/2014.

GOOGLEe. Ações manuais. Disponível em: <https://support.google.com/webmasters/answer/2604824?hl=pt-BR2604824>. Acesso em: 23/08/2014.

GOOGLEf. O que é a segmentação internacional? Disponível em: <https://support.google.com/webmasters/answer/6059209?hl=pt-BR>. Acesso em: 28/08/2014.

GOOGLEg. Status do índice. Disponível em: <https://support.google.com/webmasters/answer/2642366?hl=pt-BR2642366>. Acesso: 29/08/2014.

GOOGLEh. Busca como o Google. Disponível em: <https://support.google.com/webmasters/answer/158587?hl=pt-BR158587>. Acesso: 29/08/2014.

GOOGLEi. BLOQUEIE URLS COM O ROBOTS.TXT. Saiba mais sobre os arquivos robots.txt. Disponível em: <https://support.google.com/webmasters/answer/6062608?hl=pt-BR>. Acesso em: 29/08/2014.

GOOGLEj. Parâmetros de URL. Disponível em: <https://support.google.com/webmasters/answer/1235687?hl=pt-BR1235687>. Acesso em: 29/08/2014.

MICROSOFT. O que é malware? Disponível em: <http://www.microsoft.com/pt-br/security/resources/malware-whatis.aspx>. Acesso em: 29/08/2014.

PRIBERAM, Dicionário de Lingua Portuguesa. Blogue. Disponível em: <https://www.priberam.pt/dlpo/blogue>. Acesso em: 15/09/2014.

Anexos

Modelagem do banco de dados do Blog

Requisitos do Blog

UML do Blog