



APOSTILA BÁSICA SOBRE SEO

Fonte: Wikipédia

O QUE É SEO?

Otimização de Sites(SEO) é o conjunto de estratégias com o objetivo de potencializar e melhorar o posicionamento de um site nas páginas de resultados naturais (orgânicos) nos sites de busca.

O termo SEO (do inglês, [Search Engine Optimization](#), [Acrônimo](#) SEO) também se refere a indústria de consultoria, que trabalha na [otimização](#) de projetos e [websites](#) de seus [clientes](#). Alguns comentaristas, e também alguns profissionais de SEO, dividiram os [métodos](#) usados por seus praticantes em [categorias](#) como os "SEO de White Hat" (geralmente utilizam métodos aprovados pelos sistemas de busca, como a prática de construção de conteúdo relevante e melhoria da qualidade do site), ou "SEO de Black Hat" (utilizam truques como "cloaking", que é a camuflagem do conteúdo real da página, e [spamdexing](#)). O termo "spamdexing" (originado da fusão de [spam](#) e indexing) refere-se a prática de Spam direcionada aos motores de busca.

Os praticantes de [técnicas](#) de White Hat dizem que os Black Hat se valem de métodos que são vistos como tentativas de manipular o posicionamento dos resultados nos [Mecanismos de Busca](#) a seu próprio favor. Praticantes do Black Hat dizem que *todas* as tentativas e técnicas utilizadas para manipular o posicionamento são legítimas e o tipo de técnica utilizada para se alcançar o objetivo de melhorar o [page rank](#) é irrelevante.

Os [Mecanismos de Busca](#) indicam diferentes tipos de listagens nas [páginas de resultados](#), incluindo: [publicidade](#) paga nas [páginas de resultados](#), (SERPs - Search Engine Result Pages), conteúdo de publicidade paga por clique [pay-per-click](#), [anúncios](#), pagamento para inclusões em sistemas de busca e resultados de [busca orgânica](#).

As técnicas utilizadas no [SEO](#) são concebidas para alavancar seus objetivos no seu [website](#) melhorando o número e a posição de seus resultados nas [busca orgânica](#) para uma grande variedade de [palavras-chave](#) relevantes ao conteúdo do web site. [Estratégias](#) de SEO podem melhorar tanto o número de visitas quanto a qualidade dos visitantes, onde qualidade significa que os visitantes concluem a ação esperada pelo proprietário do site (ex. comprar, assinar, aprender algo). A Otimização para Sistemas de Busca normalmente é oferecida como um serviço profissional autônomo, ou como parte de um projeto de [Webmarketing](#) e pode ser muito eficaz quando incorporado no início do desenvolvimento do projeto de um [website](#).

Em termos de competição, um termo de busca [palavra-chave](#) com grande volume de resultados dificulta o posicionamento através programação para busca orgânica, podendo ser necessária uma ação de anúncios pagos ou links patrocinados; dependendo do número de resultados o custo de uma estratégia de marketing de internet pode ser bem expressivo. Um bom posicionamento nos resultados orgânicos de busca pode fornecer o mesmo tráfego almejado e ter um custo potencialmente mais baixo. Os proprietários de [websites](#) devem optar por otimizar seus sites para a [busca orgânica](#), se o custo do otimização for menor do que o custo da publicidade paga.

Nem todos os [sites](#) têm objetivos idênticos para o otimização em sistemas de busca. Alguns querem atingir todo o tipo de tráfego na [rede](#) e podem ser otimizados para incrementar a busca de frases comuns. Uma boa estratégia para otimização nos sistemas de busca pode funcionar perfeitamente com sites que tenham interesse em atingir um

grande público-alvo, tais como informativos periódicos, serviços de [diretórios](#), guias, ou sites que exibem publicidade com um modelo baseado em CPM (cost per mille). Muitas empresas tentam otimizar seus sites para um grande número de palavras-chave altamente específicas que indicam a disponibilidade para venda. Otimizações deste tipo, tendo em vista um amplo espectro de termos de busca, pode impedir a venda de produtos por gerar um grande volume de requisições com baixa-qualidade, com um custo financeiro alto, resultando em pequeno volume de vendas. Focalizar a qualidade do tráfego gera melhor qualidade nas vendas, permitindo que sua equipe de vendas feche negócios mais rentáveis. A otimização para sistemas de busca pode ser muito eficiente quando usada como parte de uma estratégia de [Marketing](#).

Histórico

Início dos sistemas de busca

[Webmasters](#) e responsáveis dos portais de conteúdo começaram a otimizar seus sites para os sistemas de busca em meados de -1990s, porque os primeiros motores de busca estavam iniciando a indexação do conteúdo na [World Wide Web](#). No início, os [webmasters](#) precisavam enviar o endereço do site aos vários sistemas de busca existentes na [Rede](#), para que programas como [spiders](#), pudessem "mapear" o site e armazenar as informações coletadas. O padrão e suporte das [Máquinas de Busca](#) era mapear uma página da web inteira e selecionar as chamadas palavras relacionadas na busca; então uma página com muitas palavras diferentes ampliavam a combinação de resultados, e uma página da web contendo uma relação de palavras como um dicionário teria como resultado um grande número de combinações, conseqüentemente limitando os resultados somente a nomes únicos. Os sistemas de busca então classificavam a informação por tópicos, muitas vezes exibindo como resultado das buscas páginas já expiradas, não existentes ou de conteúdo diferente ao que estava armazenado. Como o número crescente de documentos online, e vários [webmasters](#) trabalhando para aumentar o valor nos resultados em [busca orgânica](#), os sistemas de buscas mais populares começaram a classificar as páginas de resultados mais relevantes em primeiro. Iste era o início de um atrito entre Sistemas de Busca e [webmasters](#) que continua até hoje.

Os primeiros [motores de busca](#) eram orientados pelos próprios [webmasters](#). Na ocasião, as versões existentes dos [algoritmos](#) utilizados nos sistemas de busca confiaram aos webmasters e provedores de conteúdo a responsabilidade no fornecimento das informações na forma de [Categorias](#) e o uso das Meta-Tag para palavras-chave [meta tags](#) ou sistemas de busca que usavam arquivos de índice como [ALIWEB](#), fornecendo assim um guia para o [índice](#) de cada [página](#). Quando alguns [webmasters](#) começaram a abusar no uso das Meta-Tags, fazendo com que as informações de Meta-Tags das páginas não correspondessem ao conteúdo, os [motores de busca](#) abandonaram esta forma de obter informações através de Meta-Tags e desenvolveram um sistema de ranking mais complexo utilizando [algoritmos](#), elevando a filtragem das palavras e elevando o número limitado para palavras (anti-dicionário) e são mais diversas, incluindo:

- Texto dentro do Tag de título
- [Nome de Domínio](#)
- [URL](#) (de Universal Resource Locator): em português significa (Localizador Uniforme de Recursos) diretórios e nomes de arquivos
- [HTML element|HTML tags](#): cabeçalhos, negrito e textos com ênfase
- [Keyword density](#): Densidade das palavras
- Proximidade das Palavras-Chave
- [Alt attributes](#): Atributo em texto alternativo para imagens
- Textos dentro da Tag NOFRAMES

Pringle, em al. (Pringle et al., 1998) [\[4\]](#), também definiu um número de atributos dentro da codificação HTML de uma página que freqüentemente eram manipulados por provedores de conteúdo na Internet tentando melhorar a própria classificação em motores de busca. Devido a fatores que estão praticamente sob o controle exclusivo dos

responsáveis por um [site](#), os [motores de busca](#) continuaram a sofrer abusos e tentativas de classificações manipuladas. Para fornecer melhores resultados a seus usuários, os sistemas de busca tiveram que se adaptar e assegurar que as [páginas de resultados](#) mostrem sempre os resultados mais relevantes durante uma procura, ao invés de páginas inúteis cheias de palavras-chave e termos criados inescrupulosamente por webmasters na tentativa de usá-las como iscas para exibirem webpages sem link ou conteúdo. Estes fatos levaram ao nascimento de um novo tipo de motor de busca.

Sistemas de Busca Orgânica

[Google](#) foi iniciado por dois estudantes PhD da [Universidade de Stanford](#), [Sergey Brin](#) e [Larry Page](#), e trouxe um novo conceito para avaliar páginas da web. Este conceito, chamado [PageRank™](#), foi importante para o início da criação dos [algoritmos](#) do Google [2].

O PageRank™ trabalha principalmente com o fator [link](#) e usa a lógica deste sistema de ligação entre páginas como se tivesse o valor de um voto para a página em questão. O fato de uma página ser referenciada através de um link funciona de forma a validar a existência do site e dar um valor mais "digno" ao voto. O valor de cada link de referência varia diretamente baseado no PageRank da página que faz referência ao link e funciona de forma inversa no número de links de saídas de uma página.

Com ajuda de PageRank™, o Google provou ser muito eficiente em oferecer resultados relevantes nas páginas de resultados e tornou-se o sistema de busca mais bem sucedido e popular.

Pelo motivo do sistema PageRank™ fazer a avaliação de websites através de fatores externos ou fora do controle de pessoas em particular, o Google sentiu que desta maneira poderia ser mais difícil de manipular a relevância de uma página.

No entanto, webmasters já tinham desenvolvido ferramentas para manipulação de links e esquemas para influenciar o sistema de busca Inktomi. Estes métodos provaram ser igualmente aplicáveis aos [algoritmos](#) do Google. Muitos sites focalizaram suas ações em trocar, comprar, e vender links numa escala imensa. A confiança do PageRank no sistema de link como um voto de confiança para um valor da página foi subvertido como muitos webmasters procuraram vender links simplesmente para influenciar o Google em enviar mais tráfego, independentemente dos links serem realmente útil as visitantes humanos do site.

Para complicar a situação, o suporte-de-busca omitia o mapeamento de uma página web **inteira** para procurar as chamadas palavras-relacionadas nas páginas web, e uma página web contendo uma listagem do tipo dicionário ainda combinaria quase todas as buscas (exceto nomes especiais) alcançando um link-rank mais alto. Páginas de dicionários e links para esquemas podem distorcer severamente os resultados de uma página de resultados.

Já era tempo para o Google -- e outros motores de procura -- olharem para algo mais distante sobre os fatores fora de controle humano. Havia outras razões para desenvolver algoritmos mais inteligentes.

A Internet estava alcançando uma grande parcela de usuários não-técnicos, que geralmente não conheciam técnicas avançadas de pesquisa que os ajudassem a encontrar a informação que eles estavam buscando e além disso o volume e complexidade dos dados indexados tiveram um enorme crescimento em relação aos primórdios da Internet.

Para classificar os sites, os mecanismos de busca levam em consideração os seguintes aspectos:

- Idade do sítio (site)
- Há quanto tempo o domínio está registrado
- Idade do conteúdo
- Frequência do conteúdo: regularidade com a qual novo conteúdo é adicionado
- Tamanho do texto: número de palavras acima de 200-250 (não afetava o Google em 2005)
- Idade do link e reputação do sítio que o aponta
- Características padrão da página
- Pontuação negativa sobre as características da página (por exemplo, redução para sítios web com utilização extensiva de meta-tags de palavra chave, indicativos de terem sido artificialmente otimizadas)
- Originalidade do conteúdo
- Termos relevantes utilizados no conteúdo (os termos que os buscadores associam como sendo relacionados ao tema principal da página)
- Google Pagerank (usado apenas no algoritmo do Google)
- Links externos, o texto contido neste links externos e nos sítios (sites)/páginas contendo tais links
- Citações e fontes de pesquisa (indica que o conteúdo é de qualidade para pesquisa)
- Termos relacionados na base de dados do mecanismo de busca (financiar/financiamento)
- Links apontando para o site e texto-âncora destes links
- Pontuação negativa para links de chegada (provavelmente advindos de páginas de baixo valor, links de chegada recíprocos, etc.)
- Ritmo de aquisição dos links de chegada: muitos ou incremento muito rápido pode indicar atividade de comercial de compra de links
- Texto envolto aos links que apontam para fora e os links de chegada. Um link acompanhado das palavras "links patrocinados" pode ser ignorado
- Uso da tag "rel=nofollow" sugerindo aos buscadores para ignorar um link
- Profundidade do documento no site
- Métricas coletadas de outras fontes, tais como monitoramento da frequência com a qual usuários retornam clicando em voltar quando a [SERPs](#) as enviam para uma página em particular
- Métricas coletadas de fontes como [Google Toolbar](#), [Google AdWords/AdSense](#), etc.
- Métricas coletadas de compartilhamento de dados com terceiros (como provedores de dados estatísticos de programas utilizados para monitorar tráfego de sítios (sites))
- Ritmo de remoção dos links que apontam para o sítio (site)
- Uso de sub-domínios, uso de palavras-chave em sub-domínios e volume de conteúdo nos sub-domínios, com pontuação negativa para esta atividade
- Conexões semânticas dos documentos servidos
- Frequência com que os documentos são adicionados ou alterados
- IP do serviço de hospedagem e o número/qualidade dos demais sites hospedados lá
- Outras afiliações de sites com o site referenciado (eles compartilham um IP? Têm um endereço postal comum na página "Fale conosco"?)
- Assuntos técnicos como uso de 301 para redirecionar páginas movidas, mostrar um cabeçalho de erro 404 em vez 200 para páginas que não existem, uso adequado do arquivo [robots.txt](#)

- Frequência de tempo em que servidor se encontra ativo
- Se o site mostra conteúdo diferente a categorias diferentes de usuários ([cloaking](#))
- Links "quebrados" não corrigidos prontamente
- Conteúdo inseguro ou ilegal
- Qualidade da codificação HTML, presença de erros no código
- Taxa real de cliques observados pelo mecanismo de busca para as listas exibidas na [SERPs](#)
- Classificação de importância feita por humanos nas páginas com acessos mais frequentes - [ODP](#)

O relacionamento entre profissionais de SEO e as máquinas de busca

Não se encontram menções de Otimização para sistemas de busca na Usenet antes de 1997, alguns anos após a introdução dos primeiros mecanismos de busca. Os operadores de mecanismos de busca rapidamente reconheceram alguns webmasters estavam realizando esforços visando melhor classificação no ranking dos mecanismos de busca, chegando a se utilizar de técnicas de manipulação. Nos primeiros mecanismos de busca, como o [Infoseek](#), a primeira posição do ranking era obtida sem muita dificuldade, assim como era pegar o código-fonte da página mais bem colocada, colocar na própria página e enviar o endereço para instantaneamente indexar e classificar a página.

Os principais motores de busca disponibilizam guias e informações para ajudar o webmaster com a indexação e otimização do seu site: [Google](#), [Yahoo](#), e [MSN](#).

A Google tem o ~~[Ferramentas do Google para webmasters](#)~~, que ajuda a descobrir se existe algum problema de indexação do site do webmaster dentro da máquina de busca e ainda disponibiliza inúmeros dados sobre o tráfego no Google para o seu site.

O Yahoo! tem o [Site Explorer](#) (em inglês), que permite enviar sua URL gratuitamente (como também o MSN e o Google), explorar todas as páginas indexadas pelo Yahoo!, ver quais são as mais populares e disponibiliza um mapa detalhado de um site, suas páginas e links.

Estes sites de busca também permitem a criação de "adWords", que são anúncios de um site ou negócio na internet, que aparecem no resultado da busca de alguém que procura qualquer coisa relacionada com aquele produto/serviço. Uma forma que muitas empresas estão usando para fazer a divulgação e marketing na Internet. O Yahoo! tem o [Search Marketing](#)(em inglês) e a Google tem o [Google Advertising Professionals](#).

Participando dos resultados nas listagens dos sistemas de busca

Novos sites não precisam ser necessariamente "enviados" aos sistemas de busca para serem listados. Um simples link vindo de um website já estabelecido fará com que os sistemas de busca visitem o novo site e inicie o mapeamento do conteúdo. Isto pode levar alguns dias ou mesmo semanas para a indexação do link existente em um site já indexado em todos os principais sistemas de busca e spiders para iniciar a visita e indexação do novo site.

Uma vez que os sistemas de busca encontrem o site dará início então ao mapeamento das informações e páginas do site, contanto que todas as páginas usam as tags de link com padrão <a href> hyperlinks. Links para páginas que são acessíveis somente através de aplicativos em Flash ou Javascript podem não ser encontrados pelos spiders.

Os Webmasters podem instruir os spiders para não indexar certos arquivos ou diretórios através do arquivo padrão [robots.txt](#) como guia na raiz do domínio. Os protocolos atuais para as máquinas de busca e spiders exige que seja verificada a existência deste arquivo ao visitar um domínio, e mesmo que um spider ou motor de busca mantenha uma cópia oculta deste arquivo para uso na próxima visita as páginas do site, não pode atualizar essa cópia tão rapidamente quanto um webmaster . Os empresários da área de Desenvolvimento Web podem usar estes recursos para prevenir que determinadas páginas tal como carrinhos de compras ou outras páginas dinâmicas, conteúdo de uso específico de operadores apareçam em resultados sistemas de busca, assim podem manter os spider distantes de certos arquivos.

Spider de sistemas de busca podem observar um numero variado de fatores quando estiverem navegando em um site, e muitas páginas de um site podem não ser indexadas pelos sistemas de busca até que elas adquiram mais pagerank ou links ou tráfego. A distância das páginas do diretório inicial de um site pode também ser um fator decisivo se uma página vai ou não ser navegada pelo spider, bem como outras métricas importantes. Cho et al. (Cho et al., 1998) [\[3\]](#) descreveu alguns padrões para essas decisões de quando uma página visitada deve ser indexada pelo sistema de busca.

Para aqueles sistemas de busca que possuem seu próprio sistema de envio pago (como o Yahoo), pode-se poupar algum tempo ao pagar um valor de envio, no entanto, o programa de envio pago não garante a inclusão no seus resultados de busca.

Método do White Hat

White hat methods of SEO involve following the search engines' guidelines as to what is and what isn't acceptable. Their advice generally is to create content for the user, not the search engines; to make that content easily accessible to their spiders; and to not try to game their system. Often webmasters make critical mistakes when designing or setting up their websites, inadvertently "poisoning" them so that they will not rank well. White hat SEOs attempt to discover and correct mistakes, such as machine-unreadable menus, broken links, temporary redirects, or a poor navigation structure.

Because search engines are text-centric, many of the same methods that are useful for [web accessibility](#) are also advantageous for SEO. Methods are available for optimizing graphical content, including ALT attributes, and adding a text caption. Even Flash animations can be optimized by designing the page to include alternative content in case the visitor cannot read Flash.

Alguns métodos considerados próprios pelos sistemas de busca:

- Using a short, unique, and relevant title to name each page.
- Editing web pages to replace vague wording with specific terminology relevant to the subject of the page, and that the audiences that the site was developed for will expect to see on the pages, and will search with to find the site.
- Increasing the amount of original content on a site.
- Using a reasonably-sized, accurate description meta tag without excessive use of keywords, exclamation marks or off topic terms.
- Ensuring that all pages are accessible via regular links, and not only via [Java](#), [Javascript](#) or [Macromedia Flash](#) applications or [meta refresh](#) redirection; this can be done through the use of text-based links in site navigation and also via a page listing all the contents of the site (a [site map](#)).
- Allowing search engine spiders to crawl pages of your site that you want indexed without having to accept session IDs or [cookies](#).
- Developing links via natural methods: Google doesn't elaborate on this somewhat vague guideline. Dropping an email to a fellow webmaster telling him about a great article you've just posted, and requesting a link, is most likely acceptable.
- Participating in a [web ring](#) with other websites as long as the other websites are independent, share the same topic, and are of comparable quality.
- Writing useful, informational articles and offering free reprint rights for these, in exchange for a [hyperlink](#) pointing to your website.

Método Black Hat

[Spamdexing](#) é a tentativa de promoção de páginas através do uso irrelevante, principalmente comerciais, através de *técnicas decepçionantes* e que fazem uso dos [algoritmos](#) de forma abusiva. Muitos administradores de sistemas de busca consideram qualquer tipo de técnica para otimização em sistemas de e melhorar o page rank de um website como spamdexing. No entanto, em tempo um consenso comum desenvolveu na indústria o que pode ser aceitável e o que não é aceitável para reforçar uma colocação nos sistemas de busca e melhoria de tráfego resultante.

Já que mecanismos de busca operam de maneira altamente automatizada, muitas vezes webmasters usam métodos e táticas não aprovadas pelos mecanismos de busca para alcançar rankings maiores. Estes métodos por vezes passam despercebidos, a menos que um funcionário da empresa do mecanismo de busca manualmente visite o site e observe a atividade, ou uma mudança no algoritmo que controla o ranking resulte na perda dos benefícios. Às vezes as empresas contratarão um SEO para avaliar competitor's sites, e reportar métodos de otimização de resultados em mecanismos de busca antitéticos.

Spamdexing volta e meia é confundido com técnicas legítimas de otimização em mecanismos de busca, que não envolvem deceit. Spamming significa fazer sites receberem mais exposição do que eles mereceriam por suas keywords, ocasionando resultados de busca insatisfatórios. Otimização expressa o ato de fazer sites receberem o rank que eles merecem pelas mais targeted keywords, tendo por consequência resultados de busca relevantes.

Quando descoberto, mecanismos de busca podem reacionar diante daqueles utilizando métodos antiéticos de SEO. Em Fevereiro de 2006, o Google removeu tanto a BMW Germany quanto a Ricoh Germany pelo uso desse tipo de prática.[\[4\]](#)

Legal issues

A SearchKing era uma importante comunidade virtual, na qual colaboradores voluntários mantinham pequenos sítios individuais sobre assuntos de seu interesse.

O administrador da SearchKing, Bob Massa, foi um dos primeiros a enxergar que, graças à importância que o algoritmo da então crescente Google dava aos links, esses iriam adquirir um valor monetário. Massa passou a utilizar o elevado PageRank do SearchKing (então PR7) como chamariz para vender links; à época (antes de a Google criar contra-medidas que atacassem a manipulação artificial de links), um link com tal PR era capaz de melhorar o posicionamento de uma página na Google.

A Google alterou seu algoritmo, de forma a fazer que o PR da SearchKing caísse para PR2 (perdendo assim seu valor financeiro potencial). Bob Massa impetrou uma ação judicial contra a Google, alegando que a alteração ad-hoc do PageRank com o intuito de rebaixar um site específico era atitude anti-ética (unfair business practice). A Corte decidiu em favor da Google, aceitando que o PageRank é nada mais do que uma forma de a empresa expressar sua opinião sobre a relevância das páginas na internet, e como qualquer opinião ela poderia ser alterada a qualquer momento.

O caso tornou-se emblemático porque deixou claro alguns pontos: os links, que até então eram (e esse era o pressuposto do algoritmo da Google) meios de referência a páginas complementares, tornaram-se instrumentos de manipulação de rankings; vários indivíduos e empresas passaram a dar atenção específica à questão dos links (relegando um pouco a questão dos conteúdos), possibilitando o crescimento do mercado até então incipiente de Search Engine Optimization; a Google estava alerta para as mudanças, e passaria a adotar diversas alterações no algoritmo a fim de manter a qualidade de seu serviço de busca.

Qualidade e Ranqueamento de Páginas

Um [webmaster](#) que deseja maximizar o valor do seu site pode ler as diretrizes publicadas pelos mecanismos de busca, assim como as diretrizes de code publicadas pelo [World Wide Web Consortium](#). Se estas forem seguidas, e o site apresentar conteúdo atualizado, útil, original e relativamente significativo, inbound links úteis, pode ser possível alcançar uma grande quantidade de tráfego e um bom posicionamento no mecanismo de busca.

Quando um site possui conteúdo útil e cativante, há uma grande chance de que outros webmasters criem naturalmente um link para ele, aumentando o [PageRank](#) e o número de visitantes do website. Ao descobrirem um website interessante, os visitantes tendem a indicar o website, colocando como favorito ou linkando para ele, etc.

Como resultado, práticas de SEO que aumentam a qualidade do site mais provavelmente superarão táticas de manipulação do mecanismo de busca. Os melhores SEO recomendam focar-se em algo que sites de busca procuram: conteúdo relevante e útil para seus usuários.