

Sumário

SEO hoje	3
E o que eu, desenvolvedor(a), preciso saber sobre SEO?	10
Tecnicalidades do SEO	13
Conclusão	38



SEO hoje

Segundo a definição da Moz, uma das maiores empresas especializadas em Search Marketing do mundo, SEO é a **prática** de incrementar a quantidade e qualidade do tráfego de determinado website através dos **resultados de busca orgânica** (tradução livre). Existem, porém, algumas outras definições que dão mais detalhes sobre o processo e dependências de uma otimização de sites, como por exemplo a definição apresentada por Pedro Dias (ex Googler) em seu site, que diz que SEO é:

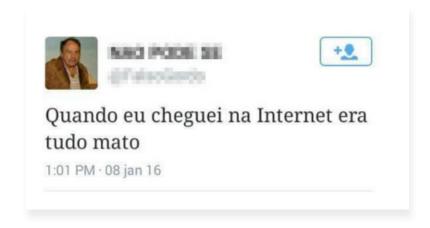
"Um conjunto de **metodologias e táticas** que, **integradas** dentro de **uma estratégia de Marketing Digital**, visam melhorias na presença digital ou o potencial de visibilidade de um site nos **resultados de busca orgânicos dos motores de busca** (como o Google ou Bing)".

Note que existem algumas diferenças entre os dois conceitos apresentados. Isso acontece sempre, independente da onde você buscou "o que é SEO". Por que? Simples: as atividades e estratégias envolvidas em um projeto de SEO estão em constante mudança. O que funcionou ontem pode não funcionar hoje e, contrariando as expectativas, voltar a funcionar amanhã. E essa é parte da magia do SEO. Nossa principal diversão é buscar fragmentos de informação que podem nos ajudar a deduzir o funcionamento de determinado fator de ranqueamento.

Um dos melhores amigos dos profissionais de SEO - e paradoxalmente o maior inimigo, por ser, na época, a representação viva do time de Qualidade de Busca (vulgo anti-spam) do Google - era o Matt Cutts. Engenheiro de software líder do time responsável por definir a qualidade dos resultados de busca do Google até janeiro de 2017, sempre dava dicas e fazia considerações bem interessantes sobre 'como melhorar o seu posicionamento nos resultados de busca'. Eventualmente ele deixava escapar alguma informação relevante para nós e, imediatamente, as redes sociais e grupos começavam a pipocar com teorias e mais teorias.

Em poucas palavras, respondendo a pergunta apresentada no título deste tópico, SEO é ciência, é métrica, é P&D. SEO é monitoramento, análise, ação e principalmente adaptação. Por fim, SEO é a sua principal estratégia para alavancar seu tráfego orgânico. Ponto.

Os primórdios dos buscadores



Lá atrás, na longínqua década de 1990, quando a internet era "tudo mato", os poucos sites que existiam só eram acessíveis se você soubesse o endereço exato deles, caso contrário, você ficaria navegando à deriva pela internet, torcendo para encontrar algo de valor a cada link maluco que você clicava. Era isso, ou ter listas e mais listas nos Favoritos do navegador ou um bloco de notas com todos os sites que te interessavam (no meu caso, como não tinha computador próprio, o bloco de notas era de papel mesmo).

Nessa época não existiam buscadores, ou, pelo menos, não como os que conhecemos hoje. Antes mesmo do surgimento da WWW (World Wide Web) já existiam alguns bots dedicados a descobrir e catalogar servidores FTP, ambientes utilizados para armazenamento e compartilhamento de arquivos, muito utilizados (ainda hoje, acredite) por instituições governamentais e de ensino. Só em meados de 1993 surgiu o primeiro crawler de fato, o Wandex. Este foi o primeiro sistema automatizado capaz de acessar e seguir links das páginas, salvando o conteúdo encontrado em seu banco de dados, nascendo assim o primeiro web crawler.

Os gigantes como Altavista, Yahoo!, ou o Cadê?, seu primo brazuca (que, apesar de terem humanos fazendo a validação de cadastro, tinha a mesma mecânica), funcionavam muito mais como um catálogo (diretórios) de sites, do que uma ferramenta de busca. Bastava submeter o site ao diretório que, com um pouco de paciência, você já poderia ser encontrado. Você informava ao diretório qual o endereço, o título, a descrição e as palavras-chave que seu

SEO hoje

site cobria e, em alguns dias (ou semanas, ou meses, variando de diretório para diretório) seu site passava a ser listado nos resultados de busca.

Naquela época praticamente não existia inteligência (humana ou artificial) por trás deste processo. Tanto é que, caso um usuário tentasse localizar seu site usando alguma palavra-chave existente em seu site, mas que por ventura você não tenha informado no seu cadastro, seu site simplesmente não aparecia. E não era só isso! A listagem dos resultados não seguia uma ordem de relevância ou qualidade, como é hoje, mas sim uma ordem alfabética. Ou seja, aparecia primeiro quem tivesse em seu nome/endereço algum número, ou a letra A, depois B... (não é a toa que a Amazon se chama Amazon).

Daí pra frente diversas empresas surgiram e desapareceram quase na mesma velocidade, algumas com métodos inovadores, outras apenas 'mais do mesmo', mas isso até 1998, o ano em que o Google começou sua história (e a mudar a forma como a busca na internet funcionava). Baseando-se na mesma proposta de crawler apresentada pelo Wandex, a verdadeira inovação do Google estava nos resultados. Pela primeira vez uma ferramenta de busca apresentava resultados realmente relevantes e mais coerentes às palavras-chave utilizadas na busca, não apenas uma lista de sites cadastrados exibida numa ordem sem um critério muito claro (ou confiável).

Provavelmente você deve estar se perguntando: como apresentar um resultado de busca altamente relevante para cada uma das (milhões de) buscas? Quais os critérios para a ordem de apresentação destes resultados? Não? Tudo bem! Sergey Brin e Larry Page, os fundadores do Google, se fizeram esta pergunta há mais de duas décadas. A resposta? PageRank.

A vida e a morte do PageRank

O PageRank é um algoritmo que mede a importância de uma página baseada na quantidade e qualidade dos links que apontam para ela. O algoritmo surgiu antes mesmo do próprio Google, sendo apresentado pela primeira vez ao público em 1996 em um artigo escrito pelos colegas Sergey Brin e Larry Page (daí vem o nome PageRank, nada a ver com 'página').

Como dito lá em cima, essa era a verdadeira inovação do mecanismo de busca do Google. Por que inovação e não invenção? Porque o algoritmo do senhor Page, além de ter sido inspirado na análise de citações publicada em 1950 por Eugene Garfield, na Universidade da Pensilvânia, também teve inspiração em um algoritmo já existente, o Hyper Search, desenvolvido por Massimo Marchiori, que mais tarde veio a fundar o - pasmem - Baidu.

Certo. Sabemos, então, que naquela época o PageRank era o responsável por definir qual resultado da busca seria apresentado primeiro. Sabemos, também, que para isso o algoritmo considera não apenas a quantidade total de links que apontam para determinada página, mas também a qualidade destes links.

Espera aí... Qualidade?

Sim! Um link nunca é apenas um link. Tem uma série de fatores que podem aumentar ou diminuir a qualidade dele. Mas, infelizmente, como quase tudo que se refere a SEO (e principalmente ao mecanismo Google), esse é um conceito extremamente subjetivo. O que sabemos, com toda a certeza, é que características como a qualidade e reputação da página de origem, a localização do link, o preenchimento dos atributos, se o link é em texto ou imagem, se tem palavra-chave exata, etc., influenciam nessa tal métrica 'qualidade'. Mas fique tranquilo, vamos nos aprofundar nisso em seguida.

Nascido o PageRank, começaram os esforços de otimização. Veja bem, não estou dizendo que o SEO não existia antes do PageRank, muito pelo contrário: o simples cadastro de um site em um diretório já poderia ser considerado SEO. Ocorre que, antes do PageRank, o posicionamento (rank) de um site em um resultado de busca ficava única e exclusivamente nas 'mãos' dos indexadores e as otimizações onpage eram basicamente relacionadas a taxa de conversão, não performance em buscas. Com o surgimento do PR, desenvolvedores e marqueteiros puderam colocar suas mãos na massa e, de fato, otimizar seus sites para um mecanismo de busca.

Mas como nem tudo na vida são flores (vide Joseph Climber), o reinado de exclusividade do PageRank não durou muito. E o motivo é, na medida em que se conhece SEO, meio óbvio: Se para aparecer antes dos meus concorrentes

nos resultados de busca, meu site precisa de mais links (e de melhor qualidade) que o site do meu concorrente, vamos sair como loucos comprando links em diversos sites. Simples, fácil e eficaz. #SQN!

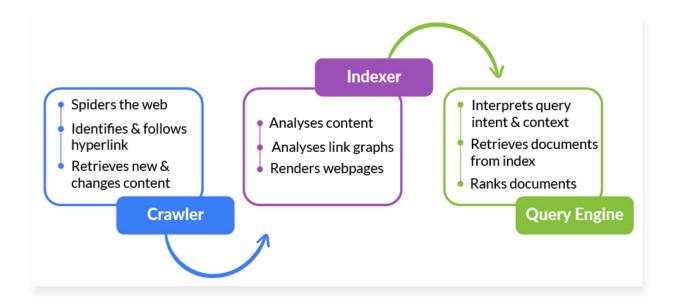
Esta estratégia acabou levando à uma onda de spam e manipulação de link tão grande que o inevitável aconteceu. Em pouco tempo o fator PageRank parou de ser o único na definição do posicionamento nos resultados, ele passou a ser uma métrica muito mais conceitual do que prática. Na medida em que o Google foi entendendo como os humanos fazem suas buscas, eles foram aperfeiçoando constantemente sua ferramenta, até se tornar o monstro que conhecemos hoje, que leva em conta mais de 200 fatores (conhecidos) para exibir um único resultado.

Quais são todos os fatores? Em que nível cada um colabora com o resultado? Infelizmente essas perguntas não podem ser respondidas, pois elas são o equivalente a receita original da Coca-Cola. Ninguém, além de um grupo extremamente seleto de pessoas, tem acesso a essa informação. Apesar de toda grande alteração no algoritmo ser amplamente divulgada e documentada (inclusive com uma previsão de impacto e comportamento esperado da SERP - Search Engine Result Page), grande parte da informação é extraoficial e normalmente as respostas às perguntas mais específicas são vagas e/ou evasivas, limitando os profissionais de SEO a supor diversas hipóteses.

Como você pode ver, não apenas o conceito do SEO tem mudado ao longo do tempo, as atividades que um profissional de SEO deve executar também mudaram. O nível de detalhe a ser observado só aumenta e, junto disso tudo, o profissional de SEO ainda precisa ser um detetive e pesquisador, captando pistas de novas alterações e testando cada uma das hipóteses.

Como funcionam os buscadores?

Como dito recentemente, não é possível ter plena certeza dos detalhes do funcionamento da ferramenta. O que sabemos - e isso é público, é que todo o processo passa por, pelo menos, três robôs diferentes, cada um com a sua função específica.



Crawler

Robô responsável por encontrar novas páginas, conteúdos e links através do processo chamado Crawling. Ao encontrar novos links em páginas já "crawleadas" o robô segue este link, lê o conteúdo e agenda o processo de Indexação. Este robô também é responsável por identificar atualização de conteúdos existentes em sites já indexados, gerando um novo processo de indexação.

Indexer

Robô responsável pela Indexação do site. Este processo é iniciado a partir da primeira visita do Crawler e consiste na leitura e análise dos conteúdos do site, perfil de links que apontam para este domínio/página e execução da renderização da página. É neste ponto que os buscadores descobrem se o seu site está ou não otimizado para dispositivos mobile, qual tecnologia que utiliza, tempo de carregamento, etc.

Query Engine

É aqui que a mágica acontece. A partir deste robô é que as análises de 'Conteúdo Buscado' X 'Conteúdo Indexado' é apresentado para o usuário, levando em consideração mais de 200 fatores.



E o que eu, desenvolvedor(a), preciso saber sobre SEO?

O escritor português Eça de Queiroz uma vez disse "O saber não ocupa espaço, e sim o vazio de nada saber". Essa máxima (que pode parecer estranha para o começo de um ebook sobre SEO) tem, na verdade, tudo a ver com a nossa realidade. Nós, desenvolvedores, designers, produtores de conteúdo e profissionais de marketing no geral, não podemos nos abster de aprender coisas novas pelo simples fato de: porque não!

Sabemos que, no mercado competitivo que temos hoje, o profissional com skills complementares (ou não) às suas "core skills" se destaca. Além disso, aprender sobre novos assuntos estimula seu cérebro a fazer novas conexões, reforçando algumas informações, alterando, ou inserindo outras, até que de repente: BUM! Você desenvolveu uma nova habilidade.



Seja pelos motivos neurológicos, seja pelos profissionais, você só tem benefícios em começar agora mesmo a desenvolver esta nova skill chamada SEO - que na verdade, convenhamos, hoje já não é apenas uma skill. Para executar um projeto de SEO a pessoa precisa ter um conjunto de diversos conhecimentos de áreas (consideravelmente) distintas entre si, como Marketing, Desenvolvimento/Infraestrutura e Produção de conteúdo.

Como nós ainda não temos um recurso para fazer o download de um conjunto de conhecimentos diretamente para o seu cérebro, a forma mais eficiente de aprender uma nova skill é lendo sobre. Por isso trouxemos este modesto guia para você, desenvolvedor, saber tudo sobre SEO. Acredito que a esta altura eu já tenha te convencido a começar a desbravar esse novo mundo chamado Otimização de Sites.

Enfim chegamos ao que realmente importa: você! Sim, você desenvolvedor ou desenvolvedora (ou aspirante a). Você é peça fundamental em qualquer projeto de SEO. Não apenas pelo fato de que, invariavelmente, precisaremos de alguém capacitado para mexer no código de um site. Também não é apenas porque vocês, mais do que qualquer outro stakeholder, têm propriedade para dizer o que pode e o que não pode ser feito. Não! Sua importância vai muito além disso.

Você, jovem coala, será a pessoa responsável pelo desenvolvimento de um dos próximos grandes sistemas ou plataformas web. Você pode ser o próximo Jobs, Gates, ou, saindo do universo do desenvolvimento, o próximo Musk. Minha missão aqui é registrar na sua cabeça que, quando for desenvolver este sistema milhonário, tenha em mente que o SEO é mais do que uma vantagem competitiva. Ele é essencial. Então, por que não te passar algumas premissas para que você possa, assim como eu estou fazendo agora, facilitar o trabalho do coleguinha?

A partir deste momento passo a relacionar os principais pontos de observação de qualquer projeto de SEO técnico, que vão desde questões relacionadas ao domínio de um site e escolha de servidores de hospedagem, até as questões relacionadas à estrutura do HTML de uma página. Deixarei as dicas mais específicas sobre conteúdo para um próximo post ou ebook.



Tecnicalidades do

SEO

Domínios, qual a melhor forma de usá-los?

Uma discussão constante entre os profissionais de SEO é o formato ideal de domínio quando o assunto é organização de conteúdos dentro de um site. Você pode ter um site em mais de um idioma, ou ter seções com conteúdos e eventualmente estrutura totalmente diferentes umas das outras. Nestes casos você pode optar por utilizar um subdomínio ou pastas dentro do seu domínio principal ou ainda utilizando domínios distintos. Mas qual a melhor opção? O que traz mais vantagem para o SEO? Antes de apresentar a resposta, é importante entender bem a anatomia de um domínio e as questões técnicas que implicam nessa escolha.

Anatomia de um Domínio

Domínios, como você já deve saber, são nomes utilizados para encontrar um determinado site na internet. Foram criados para facilitar a memorização de endereços online, que até sua criação era apenas uma sequência pouco lógica de números (IP). Uma URL de requisição é composta por pelo menos dois dos blocos apresentados abaixo, sendo o Protocolo o responsável por informar ao navegador e ao servidor que tipo de conexão que será estabelecida entre eles. Mesmo que você não digite este protocolo ao acessar uma página, ele estará lá.



A importância do SSL (HTTPS)

O termo HTTP é um acrônimo de Hypertext Transfer Protocol, que é a base de toda comunicação entre sistemas na web. O HTTPS, por sua vez, cumpre exatamente a mesma função, porém com uma camada extra de segurança baseada nos protocolos TLS/SSL.

Sempre preocupados com a experiência do usuário, o Google preza, dentre outras coisas, pela segurança de quem utiliza seu mecanismo. Por isso, em 2016, num esforço de agilizar a adoção de protocolos mais seguros por parte dos webmasters, o Google oficializou que o HTTPS passaria a ser fator de ranqueamento. Sabendo que não é bem assim pra adaptar um site inteiro para funcionar com conexão segura, deu um tempinho para que os sites se adequassem antes de reduzir a relevância deles nos resultados de busca. Esse tempinho, meus caros leitores, acabou.

Apesar de ainda ser possível ranquear bem um site sem SSL, os usuários do Chrome, que já são maioria na internet, passaram a receber alertas de páginas não segura em todas as páginas que exista algum tipo de formulário (inclusive oculto) e não tenha uma conexão SSL.

Top Level Domain (TLD)

Também conhecidas como extensões de domínio. É tudo que vem depois do '.' no final da marca/palavra-chave. Atualmente existem 7 extensões do tipo TLD disponíveis para registro, sendo que algumas possuem restrições de uso. As extensões mais utilizadas são .com, .net, .org e .gov, as duas últimas restritas à instituições governamentais.

Country Code Top Level Domain (ccTLD)

Se uma TLD é uma extensão mais abrangente de um domínio, a ccTLD é a versão específica e nacional de uma TLD. São utilizadas para segmentar sites e portais por região. No Brasil as ccTLDs mais comuns são .com.br, .net.br, .org. br, .br e .gov.br.

Generic Top Level Domain (gTLD)

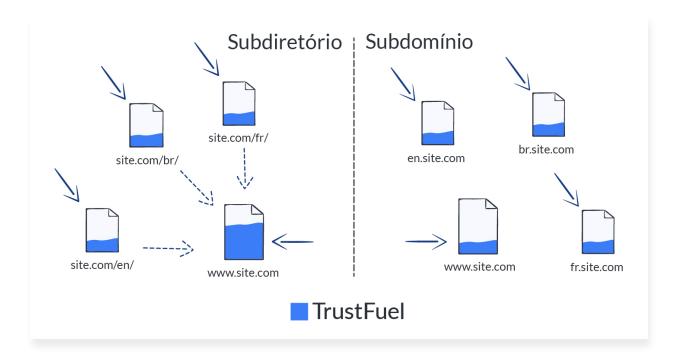
Novas extensões de domínio criadas em função da alta demanda por registros de domínios. O fato do domínio ser genérico não interfere no SEO de um site, desde que sejam respeitadas as recomendações apresentadas ao longo deste material. Alguns exemplos de domínios diferentões: .agency, .biz, .club, .io e por aí vai.

Subdomínios

Subdomínios, por sua vez, são segmentações do seu domínio principal. São aqueles caracteres alfanuméricos que precedem o domínio e que, na prática, podem ser criados sem qualquer tipo de limitação. Um subdomínio pode conter uma aplicação totalmente isolada do domínio principal, seja fazendo uso de subdiretório do mesmo host, ou até através de um apontamento de DNS, que direciona o usuário para outro Host.

O problema do Subdomínio

Os buscadores tratam subdomínios como sites diferentes e independentes do domínio principal. Isso significa que, para os buscadores, o root domain (dominio.com) e o subdomínio de site padrão (www.dominio.com) são sites diferentes. O mesmo ocorre se um determinado domínio for acessível a partir de dois protocolos distintos (http:// e https://). Ou seja, quando um link vem de um domínio e aponta para um subdiretório deste mesmo domínio (ou vice-versa) existe um repasse direto das métricas entre os conteúdos (origem → destino) e o domínio como um todo é beneficiado com os links que os subdiretórios agregam. Mas, quando um link vem de um um subdomínio, a relevância da correlação entre os conteúdos não é tão forte, transferindo menos valor para os conteúdos linkados, beneficiando os subdomínios de maneira independente.



Diretórios

Diretórios, como o nome diz, são pastas ou estruturas de pastas criadas na raiz do seu ambiente web. Esta estrutura permite melhor organização do conteúdo do site, além de proporcionar uma estrutura lógica (para humanos) da sua URL.

• •

Mas qual a melhor opção?

Bem, como tudo no mundo do SEO, depende muito de cada caso. A definição da melhor estratégia para o seu site vai depender de diversos fatores, entre eles:

- Seu site vai oferecer conteúdo em mais de um idioma?
- Terá mais de uma aplicação no mesmo ambiente?
- Que tipo de conteúdo você vai disponibilizar?
- Como os outros players fazem? Quem está na primeira posição? Como ele trata essa questão?

É importante ter em mente que, em alguns casos, é simplesmente inviável aplicar a separação dos conteúdos de uma ou outra forma. Em outros casos, por exemplo, este tipo de questionamento é totalmente desnecessário. Imagine que você está desenvolvendo um blog que vai oferecer conteúdo apenas em português e não tem intenção de utilizar o mesmo ambiente para outras aplicações, então a melhor alternativa é registrar um ccTLD e deixar tudo na raiz.

• • •

Servidores, Localização e Entregabilidade

Acredito que nesse ponto você está careca de saber que para disponibilizar um site na internet é necessário contratar uma hospedagem (aliás, você conhece a hospedagem da Umbler?). O tipo de hospedagem contratada, as características do servidor e inclusive a sua localização física interferem na performance do site. Por isso é extremamente importante que você entenda de uma vez por todas quais as diferenças entre as principais tecnologias disponíveis no mercado de Hosting.

Tecnicalidades do SEO — 17

Hospedagem Compartilhada

Normalmente a primeira opção para novos sites e aplicações. Neste tipo de servidor você tem acesso extremamente limitado às configurações do servidor, como versão do apache, iis, etc. Por se tratar de um ambiente compartilhado com outros clientes, o custo do serviço é relativamente baixo, pois uma única máquina pode comportar diversos clientes/sites, dividindo entre todos os usuários o custo de manutenção.

Prós

- Baixo Custo
- Facilidade de instalação/configuração

Contras

- Sem acesso root
- Recursos limitados
- Não é escalável
- Sem controle de hardware/software

• •

VPS (Virtual Private Server)

Em poucas palavras, o VPS é um servidor físico dividido em diversos servidores virtualizados que contam com recursos de espaço em disco, memória e processamento reservados para seu uso exclusivo. Seu gerenciamento é mais complexo do que o servidor compartilhado, normalmente ficando a cargo do cliente. Alguns hosts possuem opção de VPS gerenciado pelos profissionais de infra do hosting, o que torna a manutenção do servidor mais fácil, porém mais cara.

Prós

- Recursos exclusivos
- Acesso root
- Liberdade de configuração
- Parcialmente escalável (Memória e espaço em disco)

Contras

- Não há redundância de hardware
- Alta complexidade de configuração
- Você é responsável pelo gerenciamento do servidor
- Escalabilidade restrita aos limites físicos do servidor.
- Custo relativamente alto e normalmente envolve a contratação de um profissional especializado para fazer o gerenciamento.

Cloud

Sistema muito semelhante ao VPS, mas ao invés de ter uma máquina fracionada em diversos ambientes com recursos limitados, o Cloud é um ambiente que utiliza recursos de diversos servidores, proporcionando uma redundância de hardware, alta escalabilidade e recursos praticamente ilimitados.

Prós

- Redundância de Hardware
- Escalável
- Acesso root
- Acesso SSH
- Permite cobrança por demanda

Contras

- Gerenciamento feito pelo cliente
- Complexidade de instalação relativamente alta
- Custo mais elevado
- Normalmente serviços como Banco de Dados e e-mail não fazem parte do plano.

Localização física do servidor

Antigamente, quando o Google estava aprendendo a ser o que é hoje, ele considerava apenas a TLD para determinar a localização física de um site. Posteriormente, por volta de 2001 o Google passou a considerar não apenas o TLD, mas também o endereço IP do servidor, tornando esta variável um fator de rankeamento. Hoje em dia, com o advento dos CDNs, a localização física do servidor não afeta diretamente a posição no rank das SERPs, mas pode interferir em outros aspectos, como a experiência do usuário e o tempo de carregamento de uma página e, esses sim, são fatores de ranqueamento.

A localização física do servidor tem um peso maior quando estamos otimizando um site que tenha como foco um país específico, por isso é necessário alguns cuidados como:

- Servidor com IP dentro do range do país alvo
- ccTLD (Country Code Top Level Domain) Ex:.com.br
- Extra Tip: Reforçe o nome do país alvo no seu conteúdo.

• • •

Velocidade de Carregamento

Em estudos recentes o Google <u>identificou</u> que 40% dos seus visitantes abandonará seu site se ele não estiver totalmente carregado em, no máximo, 3 segundos. Considerando que atualmente a média de tempo de carregamento dos sites na internet é de 6 segundos, precisamos, literalmente, correr conta o tempo.

Porém, quando o assunto é velocidade de carregamento, a sigla TTFB (Time to First Byte) é a mais falada - e temida. Ela se refere ao tempo que o servidor leva para enviar o primeiro byte de resposta à uma requisição HTTP (ou HTTPS). Mas calma, pode ficar tranquilo, esta não deve ser, nem de longe, sua maior preocupação. Isso porque esta métrica depende de uma série de fatores que nem sempre estão ao alcance de um desenvolvedor ou DevOps. Por exemplo, se utilizarmos um CDN com cache, o TTFB aumenta consideravelmente, mas o tempo total de carregamento e o número de

requisições acaba sendo bem menor, compensando o 'tempo perdido' no TTFB.

São muito os fatores que influenciam no tempo de carregamento de um site. Particularmente eu costumo segmentar eles em dois grandes grupos: Fatores Externos e Fatores Internos, sendo o grupo Internos subdividido em Infraestrutura e Site.

Fatores Externos

- Largura da banda do datacenter ou servidor, caso haja alguma limitação;
- Número de saltos (redirecionamentos) necessários até o servidor ser encontrado;
- Sistema de Nome de Domínio (DNS) mal configurado ou com entradas que gerem redirecionamentos desnecessários;

• Fatores Internos (Infraestrutura)

- Recursos do Servidor: Memória, processamento, espaço e velocidade do disco, etc;
- Disponibilidade do Servidor: Diretamente relacionado ao item anterior. Quanto mais limitado for o seu ambiente, menor será a disponibilidade do site;
- Limitação de Tráfego, normalmente imposto pelo Hosting;

• Fatores Internos (Site)

- Querys SQL ineficientes;
- Funções mal escritas/executadas com loop;
- Arquivos de imagem grandes demais;
- Arquivos CSS e JS grandes de mais;
- Muitas requisições HTTP/HTTPS;
- Scripts Síncronos e/ou mal implementados.

Por sorte, a maior parte dos fatores relacionados ao tempo de carregamento de um site são internos, o que nos garante um número bem grande de frentes em que podemos atuar para reduzir o tempo de carregamento. O principal deles, na minha opinião, é implementação de um CDN, que veremos a seguir.

Tecnicalidades do SEO ________21

CDN

Se você é desenvolvedor, certamente já ouviu falar em CDN. Um CDN (Content Delivery Network ou Rede de Entrega de Conteúdo em português) é um backbone transparente que tem a função de fornecer conteúdos (páginas HTML, imagens, svg, vídeos, etc) para os usuários a partir de uma rede de servidores espalhados pelo mundo.

Utilizar um CDN torna a localização física do seu servidor uma preocupação desnecessária. Isso porque, ao acessarmos um site servido com um CDN, o download dos arquivos requisitados vai ocorrer a partir do servidor mais próximo ao usuário que está acessando. Com isso, a utilização dos CDNs entrega como vantagem a redução de latência (tempo de espera) das requisições, redução do consumo/custo de largura de banda com hospedagem e minimiza a utilização de recursos do servidor.

Um serviço de CDN pode, ou não, ter recurso de cache habilitado. Neste caso, além do usuário receber os dados a partir de um servidor próximo a ele, ele receberá uma versão cacheada do que está no servidor. O que isso significa? Um site ou aplicação ainda mais rápido. Você duvida? Vá fazer o teste. Siga o passo-a-passo deste **post**, eu espero.

• • •

Redirecionamentos e Reescritas de URL

Os redirects são, ao mesmo tempo, os maiores aliados e piores inimigos dos profissionais de SEO. Essa dicotomia acontece porque os redirects são tanto uma soluções para problemas sérios, como alteração de domínio ou mudança de URL, quanto uma dor de cabeça imensa quando mal planejados e/ou executados. Para minimizar os riscos e garantir que tudo vai funcionar direitinho, é importante ter em mente que tipo e qual o motivo dos direcionamentos.

Tipos de Redirects

Temporário (302)

Redirecionamentos temporários são utilizados nos casos em que uma página mudou de endereço (URL) ou está inacessível temporariamente, como landing pages sazonais, como de Black Friday ou cadastro de cursos. Neste caso:

- A URL original da página continua no índice dos buscadores;
- A URL original permanece com todas as métricas inalteradas, não transmitindo nada a URL temporária.
- Pode ser revertida sem prejuízos ao SEO do site.

Permanente (301)

Redirecionamentos permanentes são utilizados nos casos em que uma página saiu do ar ou foi movida permanentemente. Também utilizadas em casos de mudança de domínio e estrutura de URL. Com isso:

- A URL original sai do índice dos buscadores;
- A URL nova recebe todas as métricas da anterior;
- A reversão deste direcionamento pode afetar o posicionamento das landing pages envolvidas.

OBS: Alguns CMS possuem scripts que tratam estes redirecionamentos. Procure saber se seu CMS possui este recurso.

• •

Reescrita de URL

Diferente dos redirecionamentos, a reescrita de URL não envia o usuário para outra URL, ela apenas altera a forma como o site apresenta determinada URLs. Este recurso é utilizado para tornar URLs mais amigáveis, ocultando parâmetros ou extensões de arquivos, como .php, da sua estrutura de URL.

• • •

robots.txt, sitemap.xml e .htaccess/web.config

Estes são arquivos chave para a otimização de um site. Através deles você pode alterar algumas configurações do seu ambiente (.htaccess e web.config), permitir ou negar dos robôs acesso a alguma pasta (robots.txt) e listar, de forma hierárquica, todas as URLs de seu site. Falaremos sobre cada um deles a seguir.

robots.txt

Este arquivo é responsável por passar as orientações de como os robôs

(crawlers) devem se comportar dentro do seu site. Todos os crawlers bem intencionados são programados para respeitar estas orientação. Entre suas funcionalidades, as mais utilizadas estão:

- Permitir ou bloquear acesso a páginas ou diretórios;
- Restringir o acesso de determinado crawler a uma página ou diretório;
- Informar a URL do sitemap.xml

Confira abaixo algumas das configurações mais utilizadas

Permitir acesso de todos os robôs a todo o site

User-agent: *
Disallow:

Bloquear acesso de todos os robôs a todo o site

User-agent: *
Disallow: /

Permitir acesso apenas ao robô do google a todo o site

User-agent: Googlebot

Disallow:

Permitir acesso à algum arquivo ou subdiretório previamente bloqueado por outra regra

User-agent: *

Disallow: /imagens

Allow: /imagens/publico

Indicar aos robôs a URL do sitemap.xml

sitemap: https://seusite.com.br/sitemap.xml

Uma das dicas mais importantes que eu posso dar é: permita que o usuário responsável pelo SEO de sua aplicação tenha acesso à este arquivo. Muito frequentemente é necessário inserir alguma regra de forma manual neste arquivo e, considerando o fluxo de produção em agências, é humanamente impossível depender do desenvolvedor fazer um deploy a cada nova alteração.

Sitemap.xml

Por definição, um sitemap é o mapa do seu site e tem como objetivo facilitar a navegação dos usuários (e robôs) em seu site. O sitemap.xml, por sua vez, é um arquivo feito exclusivamente para os robôs. Sua principal função é informar aos buscadores como a hierarquia das páginas do seu site está estruturada. Sim, os buscadores podem, tranquilamente, encontrar as páginas do seu site apenas seguindo os links internos e externos, mas para que ele tenha um entendimento perfeito do seu site, é extremamente importante informá-lo quais são as páginas que ele deve visitar e como elas se relacionam entre si.

Um arquivo sitemap.xml bem feito, deve conter todas urls do site de forma hierárquica e conter informações como prioridade e data de alteração de cada uma delas. Isso é muito fácil quando temos um pequeno site institucional com poucas páginas, mas e nos casos de lojas virtuais, em que o número de URL pode facilmente ultrapassar os 5 mil? Sinto muito amigx, mas você terá que automatizar a geração desse arquivo. Para garantir que seu arquivo cumpra os requisitos e seja legível pelos robôs, seu sitemap.xml deve:

- Começar com uma tag de abertura <urlset> e terminar com uma tag de fechamento </urlset>
- Especificar o namespace (protocolo padrão) na tag <urlset>
- Incluir uma entrada <url> para cada URL, como uma tag XML pai
- Incluir uma tag filha <loc> para cada tag pai <url>

Conforme o modelo abaixo:

Apesar deste ser o formato mínimo exigido para validar um arquivo sitemap. xml, existem outras informações que podemos adicionar e que vão melhorar

ainda mais o relacionamento entre os crawlers e o seu site: as tags <lastmod>, <changefreq> e <priority>. A tag lastmod indica ao robô a data da última alteração que aquela página recebeu, muito útil para que o robô saiba se a versão que ele tem indexada é a mesma que está disponível.

A segunda tag, a changefreq, indica a frequência com que essa página é atualizada e permite ao buscador saber qual página deve priorizar ao revisitar nosso site. Imagine a página inicial de um blog, que normalmente é alterada diariamente, e uma página com formulário de contato que, provavelmente, não recebe tanta atenção. Esta tag deve trabalhar em conjunto com a tag priority, que nos permite definir uma prioridade entre as páginas de um site. A escala vai de 0 a 1 e aceita qualquer número entre estes extremos. Abaixo você confere o mesmo exemplo anterior, porém completo:

A principal limitação do sitemap.xml é o número máximo de entradas que ele aceita, que não ultrapassa os 50.000 registros. Neste casos, uma prática muito

comum que tem como finalidade organizar ainda mais a estrutura de páginas de um site, é criar diversos arquivos sitemaps, e relacionar todos em um sitemap-index.xml, que nada mais é do que um índice de todos os sitemap. xml. Essa técnica é bastante utilizada para criar um sitemap exclusivo para categorias de produtos, outro para produtos baixáveis, outro para produtos físicos, outros para... você já entendeu. Para criar o seu sitemap-index.xml basta seguir o modelo abaixo.

.htaccess / web.config

Tanto em servidores Windows quanto em servidores Linux é possível editar algumas configurações do ambiente. Para ambientes linux utilizamos o arquivo .htaccess e para Windows, o web.config. Em ambos os casos, entre outras funcionalidades, os arquivos permitem a criação de regras de redirecionamento e reescrita de URL.

Fique atento: A validação das regras nestes arquivos de configuração depende do servidor de hospedagem. No Windows, por exemplo, caso você tente alterar alguma regra que esteja bloqueada pelas configurações padrão do servidor, o sistema apresentará um erro de compilação e seu site ficará inacessível até a solução. Já no Linux a regra simplesmente não é aplicada, sem apresentação de erros no seu site.

• • •

HTML otimizado for the win!

Ter o código HTML bem otimizado é a principal premissa para o sucesso do seu projeto de SEO. É importante dominar essa linguagem de marcação, tanto quanto (ou melhor) você domina o Português. Para facilitar um pouco a sua vida, a partir desse momento darei alguns insights que, com toda certeza, irão poupar várias dores de cabeça.

Title e Meta Tags

As tags do tipo META são informações inseridas no cabeçalho (entre as tags <head> e </head>) de um documento HTML. Estas tags apresentam os detalhes sobre a informação que está disponível na página. Confira a sintaxe básica abaixo.

As tags HTML TITLE e as do tipo META possuem os conteúdos que, em princípio, o usuário do site não enxerga. São as informações contidas nestas tags que os buscadores exibem - na maioria das vezes - em suas páginas de resultado de busca.

Mas atenção: Isso não é regra! Os buscadores podem 'recortar' trechos do seu conteúdo para apresentar um resultado mais relevante ao que o usuário buscou. No exemplo acima o Google apresentou parte do conteúdo inserido na meta tag DESCRIPTION e complementou com um pedaço do conteúdo do corpo do site.

```
O melhor Cloud Host para Agências e Desenvolvedores | Umbler TITLE

https://www.umbler.com/br/desenvolvedores TURL

Umbler é um Cloud Host criado para Agências e Desenvolvedores. Temos ... Cloud com a facilidade
de gerenciamento de uma hospedagem de sites. Visualize ...

DESCRIPTION
```

Meta Informações para Social Media

Ao compartilhar um link de seu site em alguma mídia social, como Facebook ou Twitter, a própria plataforma gera um snippet contendo, normalmente: uma imagem (com base nas imagens da página de destino), o título e a descrição da página compartilhada (ambos com base nas meta tags). É possível, no entanto, definir o que será exibido nesta publicação. Isso é feito através das meta tags Open Graph (Facebook e outras redes) e Twitter Card (exclusivo do Twitter).

Open Graph

Algumas plataformas sociais, como o Facebook, utilizam estas tags para montar o que será exibido na timeline/feed. Para que uma tag OG seja válida ela deve, obrigatoriamente, conter as informações listadas abaixo, onde:

- **og:type** Representa o tipo de informação desta página (site, filme, imagem, música, artigo, quase qualquer coisa! Lista completa aqui.
- og:title Título a ser exibido no graph.
- og:image Imagem que será exibida
- **og:url** A URL destino do link.

É possível aprimorar o resultado inserindo outras tags OG como as exibidas acima. Observe que alguns TYPE possuem características exclusivas e algumas dependências. Todos os detalhes podem ser visualizados aqui: http://ogp.me/.

Observação: As tags Open Graphs são também utilizadas pelo Google para montar os resultados de busca destacados, como os que aparecem na lateral direita da SERP.

Twitter Cards

As meta tags twitter funcionam apenas na plataforma Twitter e, da mesma forma que as tags OG, existem algumas características para cada tipo de card. Veja a relação completa dos formatos possíveis em https://dev.twitter.com/docs/cards. A estrutura base está apresentada abaixo.

De acordo com a documentação do Cards, as tags podem funcionar em conjunto com as tags OG, fazendo uso das tags **og:title**, **og:description** e **og:image**. Isso significa dizer que você não precisa informar título, descrição e imagem para cada plataforma, mas como cada uma delas tem públicos (e consequentemente linguagens) diferentes, é recomendado criar os snippets de forma independentes.

A regra de ouro aqui é: Permita a edição das meta informações de suas páginas de forma independente. O administrador do site, o responsável pelo SEO e o estagiário vão amar a sua ferramenta.

Meta Robots

As Meta robots são uma forma 'in line' de passar as informações para os crawlers de como ele deve se comportar em cada página. Essas mesmas informações podem ser passadas a partir do arquivo robots.txt, mas quando se tem um site muito grande fica difícil manter o controle de todas as URLs. Sendo assim, é mais fácil utilizar as metas informações.

Esta tag aceita mais de um valor em seu content, que devem ser separado por vírgula. Ex: <meta type="robots" content="index,follow" >

Os valores mais comuns para esta meta estão listados abaixo, bem como sua respectiva função:

- index: permite a indexação desta página;
- noindex: impede a indexação desta página;
- **follow:** permite que o crawler siga os links desta página para descobrir novas;
- nofollow: impede que os links desta página sejam seguidos;
- **nosnippet:** orienta o robô a não exibir a description da página nos resultados de busca;
- noodp: orienta o Google a não utilizar a descrição do diretório DMOZ em seus resultados;
- noarchive: instrui o Google a não exibir a versão em cache da página;
- **noimageindex:** impede o crawler de indexar qualquer imagem da página.

Link

A tag HTML <link> é utilizada para fazer a relação entre o documento atual e algum documento ou recurso externo ao documento. O atributo 'rel' é o que informa o tipo de link que está sendo inserido. Esta tag é bastante utilizada para integrar arquivos .css, mas existem duas que são particularmente importantes para o SEO.

- **Canonical** Responsável por indicar aos buscadores qual a URL do conteúdo original e evitar penalização por conteúdo duplicado. Muito utilizada para content syndication e paginação de conteúdo. Todas as páginas devem ter a própria URL informada no link rel="canonical".
- Alternate + hreflang Internacionalização, Globalização, Localizations... Chame como quiser, mas não pense em ignorar esta tag se você estiver desenvolvendo um site que estará disponível em mais de um idioma. O atributo hreflang indica o idioma da versão, enquanto o href indica a URL desta versão. A detecção do idioma neste caso é feita com base no idioma do navegador do usuário.
 - Observações: As páginas devem possuir o atributo lang="xx" informado na tag HTML, na abertura do documento.
 - É possível ter versões diferentes para pt-br e pt-pt, ou en-us e en-gb, etc.

• • •

Schema.org

Schema.org é uma iniciativa colaborativa que tem como objetivo criar esquemas de dados estruturados. São um conjunto de vocabulários que desenvolvedores podem utilizar em seu código. Isso permite que os robôs entendam de forma mais assertiva o conteúdo e a semântica deste conteúdo na página.

Praticamente qualquer tipo conteúdo pode ser marcado com Schema. A lista completa e suas respectivas propriedades podem ser estudadas aqui. Alguns exemplos de tipos utilizados:

- Produtos
 - Product
 - Offer
 - AggregateOffer

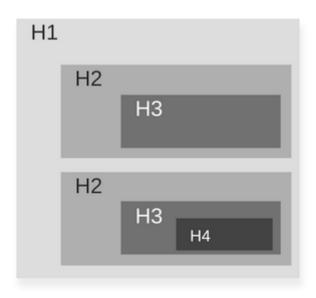
- Trabalhos Criativos
 - Movie
 - MusicRecording
 - Recipes
- Locais
 - Place
 - LocalBusiness
 - Restaurant

É importante salientar que para cada tipo de marcação existem propriedades específicas. Algumas são obrigatórias, outras opcionais. Mas o ideal é ter o máximo de informação estruturada quanto possível.

• • •

Cabeçalhos (Headers)

A organização estrutural de um conteúdo torna a leitura mais agradável e facilita a compreensão, por humanos e buscadores, sobre o conteúdo apresentado. Documentos web (Páginas HTML), assim como em um artigo científico, devem respeitar uma estrutura de títulos, subtítulos e parágrafos. Os cabeçalhos, denominados Títulos pelos editores de texto, devem ser inseridos no documento de forma ordenada e hierarquica, ou seja, só devemos inserir a tag H2 depois de já ter inserido uma tag H1.



O H1 é o único cabeçalho que deve ser usado apenas uma vez em cada documento html e seu conteúdo deve fazer referência ao tema principal da página. O WordPress (e a maior parte dos CMS) insere o título da postagem automaticamente nas páginas utilizando a tag H1, então na área de conteúdo não é recomendado inserir novamente.

Neste momento seus neurônios devem ter entrado em umas das mais acaloradas discussões do universo do SEO: Mas se eu quiser, posso usar mais de um H1? Sim, pode (mas não recomendo). Vai dar problema de validação no W3C? Se estiver usando HTML5, não. Vai impactar negativamente no meu ranqueamento? Na teoria, não. Mas então porque usar apenas um H1?

Estamos falando de uma época em que os buscadores utilizam da análise semântica de um conteúdo (além de n outros fatores) para definir se ele tem, ou não, qualidade e relevância. Por isso, costumo considerar cada página de um site ou postagem de blog, como um artigo científico que deve, obrigatoriamente, ter UM ÚNICO TÍTULO. Simples assim.

Hiperlinks

Quando falamos de hiperlinks é importante ter em mente que eles precisam ser 100% otimizados, pois são eles os responsáveis por mover os usuários (e robôs) de uma página para outra. É quesito obrigatório que todos os atributos estejam presentes, já que são as informações contidas nos atributos 'title' e 'href' e no texto âncora que informam ao usuário o que ele vai encontrar no destino do link antes dele ser clicado. A criação de links bem elaborados impacta na otimização do site como um todo.

Atributos

Obrigatórios

- href o destino do link.
- title título do link. Deve ser preenchido com uma breve descrição do conteúdo que o usuário irá encontrar na página de destino.
 Preferencialmente o título do conteúdo.

• texto âncora - Pode ser uma explicação fiel ao conteúdo do destino, a palavra chave foco do conteúdo do destino, a marca, a url, ou ainda termos genéricos como 'clique aqui'.

Opcionais

- target janela/aba alvo da ação. O correto é utilizar o _blank quando o link for externo, dessa forma mantemos a seção do usuário ativa, reduzindo bounce rate e aumentando o engajamento.
- rel quando o link for externo deve ser utilizado o valor nofollow. Este valor além de indicar que o robô não deve seguir aquele link, também impede a transferência de link juice para a página de destino.

Imagens

Assim como os links, as imagens possuem alguns atributos que devem ser utilizados para aumentar a qualidade do conteúdo no que diz respeito ao SEO. Como elementos passíveis de otimização, a imagem possui menos opções que o link, mas com igual importância.

Atributos

Obrigatórios

- nome do arquivo Utilizar palavra-chave no nome do arquivo é a forma mais eficaz de otimização de imagem. Desta forma os robôs conseguem identificar e relacionar a imagem à determinada palavra-chave mesmo que essa imagem não esteja inserida em nenhuma página html.
- alt O atributo ALT, antes de ser utilizado para SEO, é um recurso de acessibilidade. Este atributo possui o conteúdo que deve ser exibido caso a imagem não possa ser carregada. Além disso, é este o conteúdo que aparece no balão quando colocamos o ponteiro do mouse sobre a imagem e o que é 'lido' aos usuários com deficiência visual pelos assistentes de acessibilidade.
- **height e width** A otimização aqui é em relação à velocidade de carregamento da página. Ao inserir uma imagem sem os atributos de altura e largura o navegador precisa 'descobrir' as dimensões corretas dele, gerando um delay na renderização.

Minificação e Concatenação

Ter um código HTML bem estruturado, indentado e com comentários a cada linha é ótimo, mas só para desenvolvedores. Tanto os usuários quanto os robôs não dão a mínima importância para isso. Aliás, quando o assunto é SEO, manter um código HTML nessas condições vai resultar no efeito contrário. Sabemos que cada caractere (seja alfanumérico, seja especial) tem entre 1 e 4 bytes. Sabemos também que espaços vazios contam como um caractere. Sabe o nosso "Enter"? Pois então, ele também é um caractere! Um monte de bytes desnecessários. Sendo assim, quanto mais linhas, mais comentários e mais espaços em branco o código do seu documento HTML tiver, maior será o tamanho dele em bytes, consequentemente, mais tempo levará para ser baixado quando o site for acessado. Isso significa que devo criar páginas com menos HTML? Não! Quer dizer que você deve deixar esses arquivos com o menor tamanho possível. Para isso, você pode usar um recurso muito útil chamado minificação.

Atualmente existem diversas alternativas para você reduzir drasticamente o tamanho de seus arquivos HTML, JS e CSS. O principal deles, na minha opinião, é o Grunt, que possibilita, além de minimizar, concatenar vários arquivos (com a mesma extensão) em um só. Uma ótima opção quando você tem vários arquivos .js ou .css que devem ser carregado em uma página. Existem diversas outras opções online e gratuitas, o que inclui alguns plugins para WordPress, se essa for sua praia.

Dicas para criar bons Snippets

Cada um dos resultados em uma página de busca é chamado de Snippet. Eles são compostos, principalmente, pelas tags Title e Meta description inseridos no HTML da página, mas também pode apresentar trechos do próprio conteúdo da página. Para criar snippets mais ricos e que geram mais cliques, é importante atentar para alguns detalhes:

No título

- Seja específico quanto ao assunto
- Use palavras-chave foco
- Use no máximo 50 caracteres
- Faça títulos 100% únicos

Na descrição

- Descreva em poucas palavras o que o usuário vai encontrar
- Use voz ativa
- Ataque um sentimento/emoção
- Call to Action
- Use entre 130 e 150 caracteres





Conclusão

Enfim, chegamos ao final deste e-book. Se consegui fazer bem o meu papel, você chegou até este ponto com uma quantidade considerável de novos conhecimentos. Espero que tenha trazido informações úteis para o início da sua jornada no mundo da otimização de sites, e para aqueles que já desbravam este mercado, espero ter contribuído com alguns insights de como tornar o seu processo de desenvolvimento mais SEO friendly.

Certamente este material cobriu apenas uma pequena parcela de tudo que é possível falar sobre SEO, mas tomei o cuidado de trazer os pontos mais críticos para qualquer projeto. A partir daqui você está por conta própria e caberá somente a você expandir e aperfeiçoar este conhecimento. No que depender de nós, continuaremos criando e publicando mais materiais como este para ajudar você a ter ainda mais sucesso nos mercados de desenvolvimento, gestão e tecnologia.

Você tem alguma ideia para os próximos e-books? Gostaria de falar algo sobre o que leu aqui? Se sim, envie um e-mail para **gustavo@umbler.com**. Eu adoraria receber o seu feedback! Obrigado e até breve.

Conclusão — 39

Isso é a Umbler

A Umbler é uma startup de Cloud Hosting focada em agências e desenvolvedores, que oferece serviços como hospedagem de sites, registro de domínios, criação de contas de e-mail e banco de dados com pagamento pay as you go, ou por demanda. Com um painel de controle moderno e usual, o usuário consegue criar estruturas complexas de hospedagem com apenas alguns cliques. Além disso, o cliente tem total controle sobre suas aplicações, escalando, pausando ou cancelando os serviços em tempo real. No ar desde abril de 2015, conta com uma equipe com mais de 10 anos de experiência na área de hosting.

Quem é o autor?

Gustavo Sartori é formado em Marketing e pós graduando MBA em Negócios Digitais, ambos no Centro Universitário Ritter dos Reis. Tem mais de 10 anos de experiência com Web, tendo passado por algumas agências em setores como design, desenvolvimento e atendimento. Responsável pelo SEO da Umbler até início de 2018, atualmente é Product Owner do time responsável pelos websites da Umbler.

