



Universidade do Minho
Escola de Engenharia

Alexandra Diana Ccoicca Quintana

Avaliação das Técnicas de Otimização
para Motores de Busca

Avaliação das Técnicas de Otimização
para Motores de Busca

Alexandra Diana Ccoicca Quintana

UMinho | 2012

Outubro de 2012



Universidade do Minho
Escola de Engenharia

Alexandra Diana Ccoicca Quintana

Avaliação das Técnicas de Otimização
para Motores de Busca

Tese de Mestrado
Mestrado em Sistemas de Informação

Trabalho efetuado sob a orientação do
Professor Doutor Filipe de Sá-Soares

Outubro de 2012

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO,
MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Guimarães, ___/___/_____

Assinatura: _____

"Don't ever let somebody tell you that can't do something, not even me.

You got a dream, you gotta protect it.

When people can't do something themselves, they're going to tell you, that you can't do it.

You want something, go get it. Period" –

Will Smith (The Pursuit of Happiness film)

Agradecimentos

Converge neste documento uma etapa muito importante da minha vida, uma etapa inesquecível que só as pessoas que tiveram uma experiência similar podem perceber do que estou a falar. É gratificante poder culminar uma aprendizagem rica que vai contribuir para o meu futuro profissional. Deixo aqui o meu reconhecimento e agradecimento às pessoas que influenciaram e contribuíram para a finalização deste estudo.

A Deus em primeiro lugar por cuidar de mim e me encaminhar nestes quase dois anos em Portugal e por ter sempre o controlo de tudo. Agora percebo melhor que a vida sem obstáculos não é vida, e que a vida sem ele não faz sentido.

À minha linda família, aos meus pais Flora e Juan, que me inculcaram valores de amor, respeito e humildade e que sempre estiveram dispostos a ouvir-me, a guiar-me e a apoiar as minhas decisões. Aos meus irmãos Yesenia e Juan Carlos que sempre confiaram em mim e não tiveram dúvidas em me incentivar durante esta etapa. Obrigada família, amo-vos muito!

Ao meu orientador Professor Doutor Filipe de Sá-Soares, por gentilmente ter aceitado orientar-me neste tema, pela sua exemplar orientação e por ter tido muita paciência durante todo o período de realização da dissertação. Agradeço ainda, pelo conhecimento transmitido durante o meu percurso académico, nas disciplinas de Gestão da Segurança de Sistemas de Informação e Auditoria de Sistemas de Informação.

Às minhas amigas Lina, Ximena e Lola com quem tenho partilhado momentos inesquecíveis. Agradeço a sua amizade, o seu apoio, os seus conselhos e as gargalhadas, e fico muito contente pelo seu contributo para este trabalho. Obrigado por terem sido a minha família aqui em Portugal.

À Carina e à Glória do Serviço de Relações Internacionais que sempre estiveram dispostas a orientar-me e a apoiar-me durante todo este período na universidade.

Aos meus amigos de mestrado e de laboratório, especialmente Artemiso, Miguel, Quang, Rosinha, Sandra, Cris, Carina, Chetak, Zé e Clotilde pelo companheirismo, apoio e conforto que me deram ao longo de este período.

Ao Professor Doutor Carlos Sousa Pinto por ser um bom professor, uma pessoa simpática e por me ter agraciado com o seu apoio quando estava doente. Agradeço ainda pelo conhecimento transmitido durante o meu percurso académico, na disciplina de *Web Semântica*.

Às pessoas que me apoiaram nas entrevistas, pelos seus comentários, sugestões e pelo tempo que disponibilizaram para responder às perguntas, deram um grande contributo ao estudo.

Gostaria de agradecer também aos meus amigos do Grupo Bíblico Universitário que têm sido as pessoas mais simpáticas que tenho conhecido, especialmente a Sara, a Vânia e a Abigail, agradeço-lhes muito pela sua amizade, por serem pessoas fantásticas e porque sempre estiveram cientes dos meus trabalhos na universidade.

Aos meus amigos de Erasmus, não posso terminar sem os mencionar, particularmente a Nicola, Momo, Ana e Shriya, e a muitos outros que tenho conhecido durante esta etapa enriquecedora da minha vida, digo-vos que vos vou levar no meu coração.

A todos que contribuíram para o meu crescimento pessoal e profissional, muito obrigada por tudo!

Resumo

Os motores de busca são ferramentas indispensáveis na pesquisa de informação na *web*. A sua finalidade é fornecer resultados relevantes aos seus utilizadores de modo que fiquem satisfeitos. Uma das formas de o conseguirem é através da indexação, que lhes permite guardar informação útil. No entanto, devido à quantidade exorbitante de *sites* na *web*, as suas políticas de indexação mudam periodicamente. A forma de como eles operam exatamente não é conhecida e mantém-se como um segredo de negócio, uma vez que as páginas *web* listadas nos seus resultados de busca podem ficar banidas de um dia para outro. A esse processo de se posicionar nos motores de busca chama-se SEO (Search Engine Optimization), que é um conjunto de técnicas de otimização que os *sites* usam para conseguir visibilidade nos motores de busca.

O trabalho tem como objeto de estudo a avaliação das técnicas SEO correntes, *White Hat* SEO e *Black Hat* SEO. No entanto a avaliação está focada em *White Hat* SEO por estas serem as técnicas aceites pelos motores de busca. A avaliação inclui análises de três *sites* com diferentes tópicos e a sua implementação das técnicas SEO. O diagnóstico inicial foi feito mediante ferramentas para *webmasters*, que permitiram obter os fatores que impediam a visibilidade nas SERP (Search Engine Results Page). Em relação à implementação das técnicas, foram usadas ferramentas de análise, com os quais se monitorizou o comportamento de cada um deles. A parte complementar da avaliação foi a criação de um *site* (*site* de teste) baseado nas orientações dos motores de busca mais conhecidos. Uma comparação entre as técnicas SEO implementadas nos “*sites* analisados” e no “*site* de teste” determinou a eficácia das mesmas.

É de notar que uma listagem das técnicas SEO mais eficazes permite aos proprietários de *sites* criar páginas *web* de qualidade, com códigos simples, interface amigável e conteúdo otimizado, ajudando os motores de busca na indexação, e os utilizadores na busca de informação.

Palavras-chave: Otimização para motores de busca, otimização de *sites*, técnicas SEO, avaliação das técnicas SEO, *black hat* SEO, *white hat* SEO, ferramentas *analytics*, ferramentas para *webmasters*.

Abstract

Search engines are essential tools when it comes to searching for information on the web. Its purpose is to provide relevant results to its users so that they are satisfied. One of the ways to achieve it is through indexing, which allows them to store useful information. However, due to the exorbitant amount of websites on the web, indexing policies change periodically. Their exact way of operation is unknown and remains a business secret, since the web pages listed in their search results may be banned from one day to another. The process of positioning in the search engines is called SEO (English, Search Engine Optimization), which is a set of optimization techniques that websites use in order to get visibility in search engines.

The study's goal is assessing the current SEO techniques, White Hat SEO and Black Hat SEO. However the assessment is focused on White Hat SEO, due to the fact that they are accepted by search engines. The assessment includes analyzes of three sites with different topics and implementation of SEO techniques. The initial diagnosis was made through webmaster tools, which allowed getting the factors that obstructed visibility in SERP (Search Engine Results Page). Regarding the implementation of the techniques, analysis tools were used, which monitored the behavior of each of them. The complementary part of the evaluation was to create a site (test site) based on the guidelines of popularity search engines. A comparison between SEO techniques implemented in "sites analyzed" and "test site" determined their effectiveness.

It is noted that a listing of the most effective SEO techniques allows website owners to create quality websites, with simple codes, a user-friendly interface and an optimized content, benefiting the search engines in the indexing and users in finding information.

Keywords: Search Engine Optimization, SEO, SEO techniques, SEO techniques evaluation, black hat SEO, white hat SEO, analytics tools, webmaster tools

Índice

Agradecimentos	v
Resumo	ix
Abstract.....	xi
Índice	xiii
Capítulo 1 - Introdução.....	1
1.1. Contextualização.....	1
1.2. Motivação e Pertinência.....	2
1.3. Objetivos da Investigação	3
1.4. Abordagem de Investigação.....	4
1.5. Estrutura do Documento	9
Capítulo 2 - Revisão do Estado de Arte	11
2.1. A Internet e a Web	11
2.2. Motores de Busca.....	12
2.3. Search Engine Optimization (SEO).....	15
2.4. Técnicas SEO.....	20
2.4.1. Técnicas <i>White Hat</i> SEO.....	20
2.4.2. Técnicas <i>Black Hat</i> SEO	28
2.5. Ferramentas de <i>Web Analytics</i>	31
2.6. Ferramenta de Apoio para os <i>Webmasters</i>	35
2.6.1. <i>Google Webmaster Tools</i>	35
2.6.2. <i>Bing Webmaster Tools</i>	36
2.7. Considerações Finais	36
Capítulo 3 - Descrição do Trabalho Efetuado	39
3.1. Visão Global do Trabalho.....	39
3.1.1. Recolha de Dados.....	40
3.1.2. Implementação	41
3.1.3. Teste	51
3.3. Entrevistas.....	58
Capítulo 4 - Resultados	61
4.1. Resultados da Entrevista.....	61
4.2. Resultados da Avaliação	64

4.3. Comparação das Técnicas SEO	80
Capítulo 5 - Discussão.....	83
5.1. Avaliação das Técnicas.....	83
5.2. Resultados nos Motores de Busca	84
5.3. Visibilidade.....	86
5.3. Google e SEO	88
Capítulo 6 - Conclusões.....	89
6.1. Conclusões do Estudo.....	89
6.1. Limitações.....	91
6.2. Contribuições.....	91
6.4. Recomendações e Trabalhos Futuros	92
6.5. Considerações Finais	92
Referências	93
Anexos.....	99
Anexo I – Mensagem Eletrónica de Pedido de Entrevista.....	99
Anexo II – Guião da Entrevista	100
Anexo III – Resumo das Entrevistas.....	101
Apêndice.....	105
Apêndice I – Diretrizes dos Motores de Busca mais Conhecidos	105

Índice de Ilustrações

Ilustração 1. Estrutura da Pesquisa de Dissertação.....	4
Ilustração 2. Análise do Tema de Dissertação.....	6
Ilustração 3. Crescimento da Internet nos Últimos Anos	11
Ilustração 4. Motores de Busca Baseados em <i>Crawler</i>	14
Ilustração 5. Resultados Orgânicos e Pagos nos Motores de Busca.....	17
Ilustração 6. Processo SEO.....	20
Ilustração 7. Processo do Estudo Efetuado.....	40
Ilustração 8. Alterando os Tags de <i>Site A</i>	50
Ilustração 9. <i>Mock-up</i> do <i>Site</i> de Teste.....	53
Ilustração 10. Tendência de Busca a Nível Mundial para as Áreas de Negócio, Educação e Viagens no Ano 2012.	54
Ilustração 11. Tendência de Busca a Nível Mundial para o Termo <i>Erasmus</i> no Ano 2012	55
Ilustração 12. Escolha do Nome de Domínio do <i>Site</i> de Prova.....	57
Ilustração 13. Processo de Avaliação das Técnicas SEO	64
Ilustração 14. Evolução do <i>Site A</i> quanto o Tráfego no <i>Google Webmaster Tool</i>	67
Ilustração 15. Evolução do <i>Site B</i> quanto o Tráfego no <i>Google Webmaster Tool</i>	68
Ilustração 16. Evolução do <i>Site C</i> quanto o Tráfego no <i>Google Webmaster Tool</i>	69
Ilustração 17. Evolução do <i>Site A</i> quanto o Tráfego no <i>Bing Webmaster Tool</i>	70
Ilustração 18. Informação Estatística do Tráfego Gerado no <i>Site A</i>	71
Ilustração 19. Informação Estatística do Tráfego Gerado no <i>Site B</i>	72
Ilustração 20. Informação Estatística do Tráfego Gerado no <i>Site C</i>	72
Ilustração 21. Página de Entrada do <i>Site</i> de Teste	74
Ilustração 22. Página Utilizador do <i>Site</i> de Teste.....	75
Ilustração 23. Evolução do <i>Site</i> de Teste desde a sua Criação	76
Ilustração 24. Informação Estatística do Tráfego Gerado no <i>Site</i> de Teste.....	78
Ilustração 25. Posição do <i>Site</i> de Teste no Google.....	79
Ilustração 26. Posição do <i>Site</i> de Teste no Bing.....	79
Ilustração 27. Posição do <i>Site</i> de Teste no Yahoo.....	80

Índice de Tabelas

Tabela 1. Ferramentas Estatísticas mais Populares.	33
Tabela 2. Conjunto de Técnicas SEO	41
Tabela 3. Descrição dos Peritos Entrevistados	59
Tabela 4. Período da Implementação das Técnicas SEO nos <i>Sites</i> Analisados	66
Tabela 5. Resultados das Técnicas SEO em Base as Métricas Obtidas da Análise dos <i>Sites</i> A, B e C.....	73
Tabela 6. Período da Implementação das Técnicas SEO no <i>Site</i> de Teste.....	75
Tabela 7. Resultados das Técnicas SEO no <i>Site</i> de Prova.....	77
Tabela 8. Técnicas SEO Correntes mais Eficazes	81

Índice de Relatórios

Relatório 1. Visão Global do <i>Site A</i>	44
Relatório 2. Visão global do <i>Site B</i>	45
Relatório 3. Visão Global do <i>Site C</i>	46
Relatório 4. Estado Atual do <i>Site A</i>	48
Relatório 5. Estado Atual do <i>Site B</i>	49
Relatório 6. Estado Atual do <i>Site C</i>	50

Lista de Abreviaturas

CSS – Cascading Style Sheets

CPC – Cost Per Click

FTP – File Transfer Protocol

HTML – HyperText Markup Language

IP – Internet Protocol

ITU – International Telecommunications Union

NDTL – Networked Digital Library of Theses and Dissertations

PPC – Pay Per Click

RCAAP – Repositório Científico de Acesso Aberto de Portugal

RSS – Really Simple Syndication

SDUM – Serviços de Documentação da Universidade do Minho

SEM – Search Engine Marketing

SEO – Search Engine Optimization

SERP – Search Engine Results Pages

SMO – Search Marketing Optimization

SMM – Social Media Marketing

TCP – Transmission Control Protocol

URL – Uniform Resource Locator

XML – eXtensible Markup Language

VERONICA – Very Easy Rodent-Oriented Net-wide Index to Computerized Archives

WWW – World Wide Web

Capítulo 1 - Introdução

1.1. Contextualização

A fim de indexar¹ toda a informação contida na *Web*, informação que tem vindo aumentar nos últimos anos, criaram-se os motores de busca. Um dos pioneiros foi Archie [Emtage e Deutsch 1992] que permitiu aos utilizadores² procurar ficheiros específicos localizados em *sites* públicos de FTP (File Transfer Protocol). Posteriormente surgiram Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives), Lycos, Excite, entre outros [Chu e Rosenthal 1996] até aos mais conhecidos de hoje: Yahoo, Google e Bing. Assim, já há mais de uma década que os *Webmasters*³ tentam otimizar⁴ *sites* para os motores de busca.

Ao processo de otimização de *sites*, seja no seu *design* ou no seu conteúdo, chama-se SEO (Search Engine Optimization). Atualmente SEO faz parte da estratégia *Marketing Online*⁵ [Chaffey 2008; Mitchell 2012; Wynne 2012], sendo que algumas das suas técnicas são: submissão a diretorias, *linkBaiting*⁶, criação de comunidades, uso de *sitemap* XML (eXtensible Markup Language), entre outras. A definição do SEO é diversa, há quem o considere como um conjunto de técnicas [Gandour e Regolini 2011], alguns como uma prática [McVittie 2012], outros como um processo [Jones 2008], uns como uma ciência [Ledford 2009] e ainda, outros como uma arte [Lieb 2009] que tem por objetivo melhorar o volume e qualidade de tráfego⁷ nos *sites*, através de resultados de busca natural (orgânica ou algorítmica) [Lieberam-Schmidt 2010]. Para efeitos deste estudo, SEO é considerado como uma parte complementar da estratégia *Marketing Online* que possui um conjunto de técnicas para otimizar *sites*. A maioria das atividades SEO aspira por elevadas posições nas páginas de resultados dos motores de busca, dado as taxas de clique serem usualmente mais elevadas nas posições superiores [Pizam 2010].

¹ Organizar informação na *web*, ou seja colocá-lo num índice. Para ter uma ideia clara compara-se com o índice de um livro.

² Ao longo de todo o documento, utilizador ou visitante é tratado da mesma forma.

³ Chama-se também *Web Designer*, Proprietário de *site*, Responsável do *site*, ou Dono do *site*. É a pessoa responsável por uma ou mais páginas.

⁴ Dar o valor máximo a um processo, produto ou serviço.

⁵ Conhecido também como *Marketing Digital*, *Internet Marketing* e *Web Marketing* é tudo sobre o *marketing* para atuais e potenciais clientes através da Internet [Miller 2012].

⁶ Para mais detalhes ver a seção 2.4.1.

⁷ Quantidade de dados enviados e recebidos pelos visitantes para um *site*.

Os resultados da aplicação das técnicas SEO nos *sites* geralmente variam e alguns levam mais tempo do que outros. Uma das causas é a grande quantidade de *sites* na *internet*. No entanto, existem outros fatores, como o conteúdo, as etiquetas de título e até mesmo o código do *site*. O que normalmente é feito neste caso é a procura de palavras-chave ótimas que possam ser usadas no conteúdo dos *sites*, de forma que os motores de busca classifiquem estes *sites* com base nessas palavras-chave [Schimmel 2010], e se posicionar nas SERPs (Search Engine Results Pages).

Em todo este processo de otimização de *sites*, as técnicas SEO, chamadas também Técnicas *White Hat* SEO [McDonald e Zayat 2009], por vezes não são usadas de maneira correta, pelo que muitos *Webmasters* incorrem em práticas proibidas pelos motores de busca, conhecidas habitualmente como Técnicas *Black Hat* SEO [Malaga 2008]. A avaliação da eficácia das técnicas SEO é uma forma de perceber como estão a ser usadas as técnicas, no sentido de verificar os fatores que impedem a visibilidade dos *sites* nos motores de busca, e se elas realmente estão a cumprir com as boas práticas para otimizar os *sites*. Avaliações deste tipo trazem benefícios substanciais tanto para os motores de busca, como para os *sites*, porque quebram a barreira existente entre ambos, o que significa permitir aos motores de busca criar novas funcionalidades de indexação, e aos *sites* otimizar de maneira correta, desde a sua criação até o seu lançamento. Uma meta comum que ambos visam é satisfazer os utilizadores, pelo que uma informação de qualidade, em tempo real e adaptadas as necessidades do utilizador é fundamental.

1.2. Motivação e Pertinência

A *Web* tem-se convertido num serviço indispensável na vida das pessoas, qualquer informação que se queira encontrar, seja profissional ou pessoal, pode estar contida nela. De modo que, viver nestes tempos sem uma conexão à *Internet* é quase impossível. Uma das ferramentas mais usadas na *Web* para busca de informação são os motores de busca, que visam retornar informação valiosa aos utilizadores. Toda esta informação adaptada para a *Web* está estruturada em documentos HTML (*HyperText Markup Language*) chamados páginas. O conjunto destas páginas origina um *site* que pode ser visualizado pelo *browser*. Um *site* basicamente é um cartão de visitas digital, para empresas, organizações e pessoas.

As empresas de hoje focam-se cada vez mais nos conteúdos *online* para publicitar os seus produtos e serviços devido a um grande número de indivíduos estarem conectadas à internet. Serviços como *e-mail*, motores de busca, redes sociais, *blogs*, entre outros, são usados para alcançar, informar e conectar estas pessoas. Deste modo, uma estratégia de *Marketing Online* é recomendável para qualquer tipo de negócio.

Uma das motivações fundamentais deste estudo é o aumento do uso das técnicas SEO nos últimos anos, como parte da estratégia comercial das empresas. A eficácia com que operam os motores de busca e a informação quase personalizada que retorna aos utilizadores tem contribuído para este crescimento. No entanto, o esforço que se realiza com as técnicas SEO, muitas vezes é manipulado de maneira imprópria, devido a concorrência de milhões de *sites* existentes na *web* com a mesma temática e a vontade de querer impulsionar maiores receitas dos seus negócios. O motivo pelo qual, surgiram algumas questões: Se for possível avaliar as técnicas SEO, até que ponto as empresas ou as pessoas que trabalham nesta área estão a utilizar as técnicas corretamente? Serão estas técnicas eficazes, ou não? Sendo assim, as diretrizes providenciadas pelos motores de busca, a base principal para avaliar as técnicas SEO. Por conseguinte, a relevância deste estudo centra-se na avaliação da eficácia das técnicas SEO sobre *sites* com diferentes tópicos.

1.3. Objetivos da Investigação

O objetivo principal deste estudo está focado na seguinte questão de investigação: Atualmente, quais as técnicas SEO mais eficazes? Para responder a esta questão, estabeleceram-se vários objetivos que visaram orientar a investigação.

O primeiro objetivo foi identificar e determinar as principais características das técnicas SEO através da revisão de literatura. A revisão foi ressaltada no funcionamento dos motores de busca, as técnicas SEO mais correntes e as ferramentas de suporte SEO.

O segundo propósito deste estudo foi diagnosticar *sites* na vertente de SEO, de modo a conhecer os factores que impedem a visibilidade nos motores de busca e implementar as técnicas SEO sobre os resultados obtidos.

O terceiro objetivo deste estudo foi criar um *site* de acordo com as sugestões encontradas nas diretrizes dos motores de busca, de modo a analisar a evolução das técnicas SEO implementadas.

Por último, pretendeu-se validar as técnicas SEO com base as métricas encontradas durante a revisão do estado de arte. A validação fez-se através de uma comparação entre os resultados obtidos do diagnóstico dos *sites* com os do *site* criado.

O alcance dos objetivos identificados permitiu a construção de uma tabela, com as técnicas SEO mais eficazes, ordenadas pela sua importância.

1.4. Abordagem de Investigação

Na seção anterior foram definidos os objetivos que guiaram este trabalho de dissertação. Nesta seção apresenta-se a abordagem de investigação focado na abordagem metodológica e nas técnicas utilizadas durante o estudo. Subsequentemente apresenta-se, o modelo da estrutura da pesquisa da dissertação (Ilustração 1).

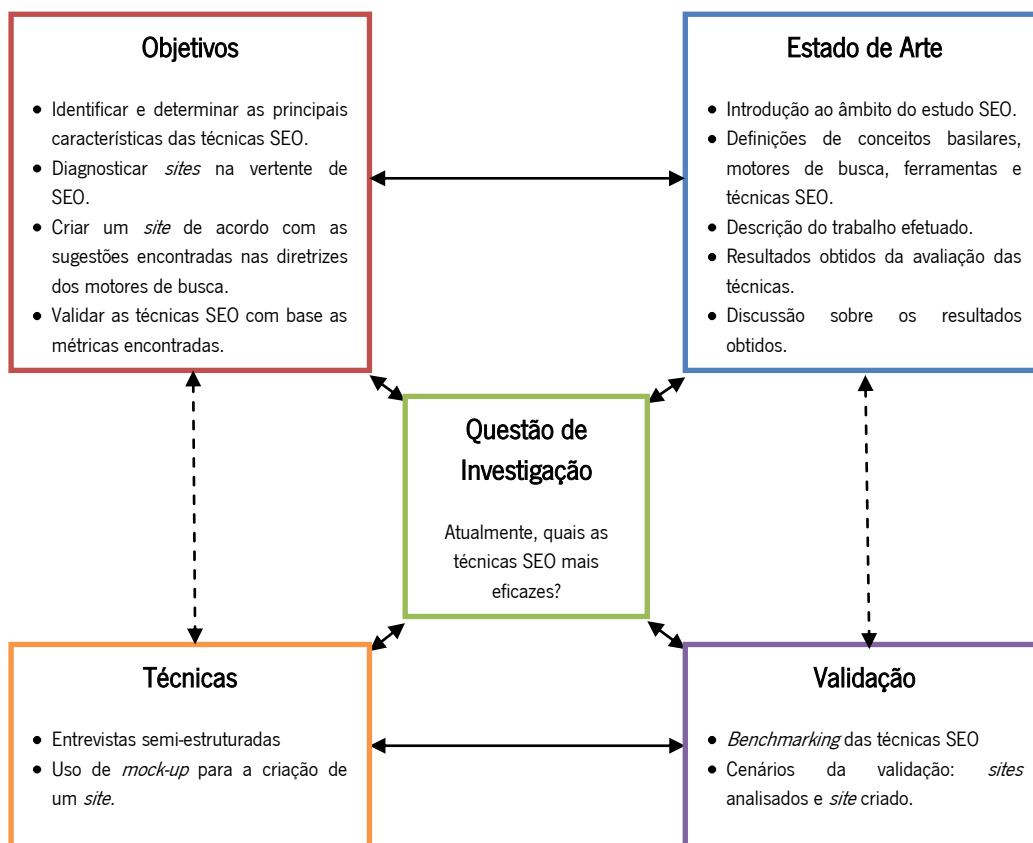


Ilustração 1. Estrutura da Pesquisa de Dissertação. Adaptado de Maxwell [2004].

Quer seja uma dissertação de mestrado ou quer seja uma tese de doutoramento é essencial utilizar uma abordagem metodológica adequada por forma a atingir os objetivos enunciados. Neste sentido a abordagem metodológica utilizada foi mista (Quantitativa e Qualitativa), uma vez que se pretendeu conhecer a eficácia das técnicas SEO através de um seguimento do progresso de cada um delas em diferentes *sites* e se realizou entrevistas qualitativas de forma a prover elementos adicionais à análise quantitativa. A pesquisa quantitativa positiva permite aos investigadores de TI responder a questões de investigação sobre a interação entre as pessoas e os computadores [Straub et al. 2004], basicamente está baseado na medição da quantidade ou valor [Kothari 2009] do fenómeno que está a ser estudado. Os dados podem ser recolhidos de ficheiros ou entrevistas e as quantidades podem ser medidas por ferramentas ou pacotes estatísticos [Straub et al. 2004]. Entre os métodos⁸ que a abordagem apresenta, o *Field Experiment* é o mais apropriado para este estudo, uma vez que, o estudo foi realizado num ambiente real onde os utilizadores interagiram.

De forma a validar as soluções encontradas no *Field Experiment* fez-se uso do modelo *Benchmarking*. O *Benchmarking* é usado para mostrar que uma solução tem um desempenho razoável, ou é melhor do que qualquer outra ou outras soluções disponíveis [Vaishnavi e Kuechler, 2004]. Esta definição encaixa perfeitamente no propósito do estudo que é conhecer a eficácia das técnicas SEO. Em relação ao cenário de teste, Vaishnavi e Kuechler [2004] continuam afirmar que se o *benchmark* não está disponível existe a possibilidade de criar um cenário de teste ou uma classe de cenários para avaliar uma solução, bem como qualquer outra solução disponível. Para este propósito, o cenário criado foi o *site* de teste, onde se avaliou a eficácia de cada técnica SEO requerida ou uma combinação delas. Da mesma forma, se fez uso de cenários existentes para avaliar as técnicas SEO, neste caso de três *sites* com diferentes tópicos fornecidos por pessoas externas ao estudo. Estas soluções foram comparadas com as outras soluções a fim de validar a eficácia de cada uma das técnicas.

Por outro lado, de forma a sustentar e complementar o conhecimento da autora do presente estudo, realizou-se uma revisão bibliográfica. A revisão bibliográfica foi desenvolvida com base em material já elaborado ou extraídos de fontes alternativas, principalmente livros e artigos científicos, que permitiram ao investigador ter um foco

⁸ *Field experiment, Lab experiment, Free simulation experiment, Experimental simulation, Adaptive experiment, Field study, Opinion research/ surveys e Archival research* são alguns dos métodos que a pesquisa quantitativa positiva apresenta.

maior no seu trabalho do que obtê-lo diretamente. Para este fim, recorreu-se a bases de dados bibliográficas, revistas científicas, atas de conferências e a outros conteúdos necessários no desenvolvimento do trabalho. Os repositórios digitais usados para a pesquisa bibliográfica foram: 1) SDUM (Serviços de Documentação da Universidade do Minho), 2) *Google Scholar*, 3) *Web of Knowledge*, 4) *Scopus*, 5) NDTL (Networked Digital Library of Theses and Dissertations), 6) DRIVER e 7) RCAAP (Repositório Científico de Acesso Aberto de Portugal).

Antes de formular as palavras-chave para a pesquisa bibliográfica foi feita uma análise do tema da investigação, separando as palavras umas das outras com a finalidade de conhecer o que se tinha de fazer, o que limitava o tema e o que este abrangia (Ilustração 2).

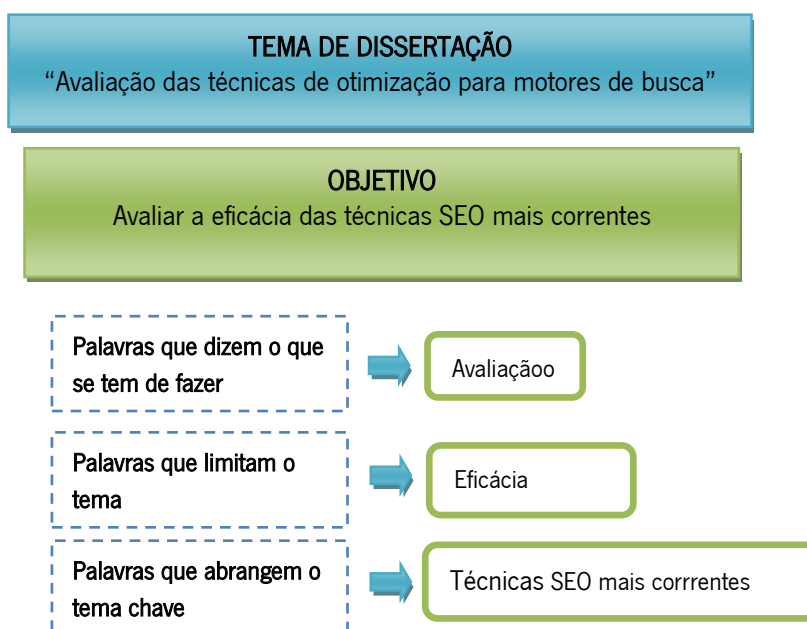


Ilustração 2. Análise do Tema de Dissertação.

Na pesquisa efetuada utilizaram-se palavras-chave referentes aos motores de busca, técnicas de otimização e ferramentas analíticas de páginas *web*, maioritariamente em Inglês, e um grupo mais pequeno em Português e Espanhol. Para completar, foi realizada a revisão bibliográfica sobre *marketing digital*, de forma a suportar o tema de investigação.

No que respeita à recolha de dados, existe uma variedade de técnicas como os relatórios, a observação do participante⁹, grelha de reportório¹⁰, o incidente crítico¹¹, *laddering*¹², e os mais comuns, questionários e entrevistas [Rugg e Petre 2006]. A técnica que mais se enquadrou no estudo foram as entrevistas, uma vez que se pretendeu obter sugestões, comentários e uma impressão geral sobre o tema em discussão a partir da perspectiva de peritos SEO na atualidade. Na maioria de textos as entrevistas estão divididas em três tipos: estruturada, semiestruturada e não estruturada. As entrevistas estruturadas possuem uma estrutura definida antes de ser realizadas, esta pode ser uma lista de temas, ou uma lista de questões elaboradas em um guião. É praticamente um questionário falado. As entrevistas não estruturadas, como o seu nome indicada, não precisam de estruturas pré-determinadas para serem realizadas. O investigador utiliza a sua habilidade e os seus próprios critérios para prosseguir com a entrevista. Quanto as entrevistas semiestruturadas, possuem alguns tópicos ou questões pré-determinadas, no entanto são suficientemente flexíveis para introduzir tópicos interessantes quando eles surgem durante a interação entre o investigador e o entrevistado [Rugg e Petre 2006].

Dependendo do tipo de entrevista, os participantes têm mais ou menos a oportunidade de responder a perguntas nos seus próprios termos. Tradicionalmente, as entrevistas, independentemente do seu propósito, eram realizadas cara a cara mas hoje, as tecnologias de comunicações eletrônicas têm permitido aos investigadores utilizar novas abordagens, pelo que podem ser dirigidas *online*. As entrevistas *online* são geralmente realizados por *e-mail* ou por *chatting*, usando um software em tempo real [Mann e Stewart 2000]. As vantagens que apresentam são [Mann e Stewart 2000; Wohlin et al. 2003]:

- O entrevistador geralmente diminui o número de “não sei” e “sem resposta”, porque o participante pode responder as perguntas sobre o questionário.
- Amplia o acesso aos participantes. Como a Internet tem um acesso geográfico amplo, permite aos investigadores acederem a populações de difícil acesso.
- Menor custo e poupança de tempo.

⁹ Técnica usada nas pesquisas para observar o comportamento do participante.

¹⁰ O procedimento geral da grelha de repertório consiste numa forma de entrevista estruturada e focalizada num determinado domínio da experiência do sujeito [Fernandes 2001].

¹¹ Consiste em um conjunto de procedimentos para a coleta de dados em estudos de pesquisa onde está envolvida a atividade humana [Kremer 1980].

¹² É uma técnica de coleta de dados onde o pesquisador usa um pequeno conjunto de questões para esvaziar sistematicamente o conhecimento dos participantes.

- Não é necessário organizar equipamentos de gravação e custeio para a transcrição.
- Manipulação simples de dados.
- Oferece um ambiente seguro, privado e familiar.

Se o investigador está pensar utilizar esta técnica, tem ter em consideração que a falta de conhecimentos básicos de informática, a insuficiente formação para moderar grupos e as dificuldades em contactar e recrutar participantes, podem ser desvantagens quando se recorre a este tipo de técnicas [Mann e Stewart 2000].

Segundo os parágrafos anteriores, o tipo de entrevista aplicado para este estudo foi o semi-estruturado uma vez que se levou em consideração um guião pré-determinado (Anexo I) e realizou-se de maneira *online* a fim de aprofundar os temas tratados no estado de arte. As questões focaram-se nas técnicas SEO, as ferramentas SEO e algumas referências dos motores de busca mais conhecidos. As características da entrevista permitiram fazer outras perguntas na tentativa de compreender melhor a informação obtida e facilitou a adaptação do entrevistado. As questões foram semelhantes para todos os participantes mas não se seguiram uma ordem restrita. O propósito das entrevistas foi obter informação global do tema de investigação e procurar opiniões de peritos SEO de modo a atingir os objetivos pretendidos da pesquisa. Uma das principais razões para o uso de este tipo de entrevistas foi reduzir o tempo e o custo que implicava fazer todo o processo, e adaptar-se facilmente à disponibilidade dos participantes. O processo da entrevista é descrito no capítulo 3 e os resultados são apresentados no capítulo 4.

Igualmente fez-se uso de um *mock-up* para criar o *site* de teste, mas não antes de fazer um estudo de mercado das características que iria a possuir, o tema a tratar e o que se espera dele. [Miller, 2000] sugere uma estrutura de qualidade de um *site*, que foi considerado como base para a realização do *mock-up*. Mais detalhes sobre este tópico são apresentados na secção 3.2 do capítulo 3.

1.5. Estrutura do Documento

A presente dissertação é constituída por seis capítulos, conforme referido nos parágrafos que se seguem.

No primeiro capítulo é feita uma introdução ao âmbito do estudo, de forma que o leitor se enquadre na área a analisar e crie uma base de raciocínio para os seguintes capítulos. Neste sentido, o capítulo em questão compreende uma breve descrição do enquadramento do estudo, bem como os motivos que levaram ao aprofundamento da área de investigação. São igualmente apresentados os objetivos da investigação assim como, a abordagem metodológica e a seção que vem sendo descrita.

No segundo capítulo apresentam-se as definições dos conceitos basilares deste estudo, e procede-se à revisão da literatura no que se refere aos motores de busca e à forma como estes funcionam, quais são os mais populares, e as diretrizes que apresentam cada uma delas. Também são abordadas as técnicas SEO, com as perspectivas de diferentes autores e são divididas em dois grupos, *White Hat* SEO e *Black Hat* SEO. São ainda mencionadas as ferramentas *analytics* e para *webmasters* mais usadas que ajudam a ter o controlo dos sites enquanto são otimizados.

No terceiro capítulo reporta-se a descrição do trabalho efetuado, descrevendo-se todos os processos que contribuíram para a avaliação das técnicas SEO. O trabalho foi dividido em três grandes processos, dentro dos quais surgem outros subitens que fazem parte do trabalho e que também são descritos.

No quarto capítulo são apresentados os resultados obtidos na avaliação das técnicas SEO implementadas. São mencionados os resultados das entrevistas, os resultados das avaliações de cada *site* com as respetivas ilustrações e duas tabelas com o resumo das técnicas implementadas nos *sites* analisados e no *site* de teste.

No quinto capítulo, discutem-se os resultados obtidos na avaliação das técnicas SEO. As discussões são divididas por secções apresentando-se em primeiro lugar as discussões das avaliações, e em segundo os motores de busca e dos factores que impedem os *sites* conseguir visibilidade nas SERP. Ainda é discutido como o SEO é observado pelo Google.

Por fim, no sexto capítulo, conclui-se a dissertação, apresentando-se as conclusões mais importantes, e são apresentadas as contribuições, limitações, e recomendações para trabalhos futuros.

Capítulo 2 - Revisão do Estado de Arte

2.1. A Internet e a Web

Se prestarmos atenção, o termo *web* tem sido substituído pelo termo *internet*, e é comum ver em bibliografias ou ouvir pessoas falar sobre *internet* quando se estão a referir à *web*. As definições de [Costa et al. 1995] podem esclarecer as diferenças: *internet* refere-se a uma rede de computadores que abarca praticamente todo o globo terrestre e a *web* refere-se a um sistema de informação hipertexto¹³ e hipermedia¹⁴, distribuído, cooperativo e heterogéneo. Portanto, o termo *internet* não é sinónimo da *web*, a *internet* é usada pela *web* para transmitir informação entre computadores.

O uso da Internet nos últimos anos tem crescido consideravelmente. De acordo com as estatísticas feitas pela ITU (International Telecommunications Union), o uso da internet nos últimos cinco anos, desde 2006 até 2011 cresceu 17% tendo em conta a população mundial. As estatísticas apresentadas para o ano 2006 foram de 18 % e para o ano 2011 foram de 35%, portanto hoje, um terço da população mundial encontra-se ligado à internet (ver Ilustração 3).

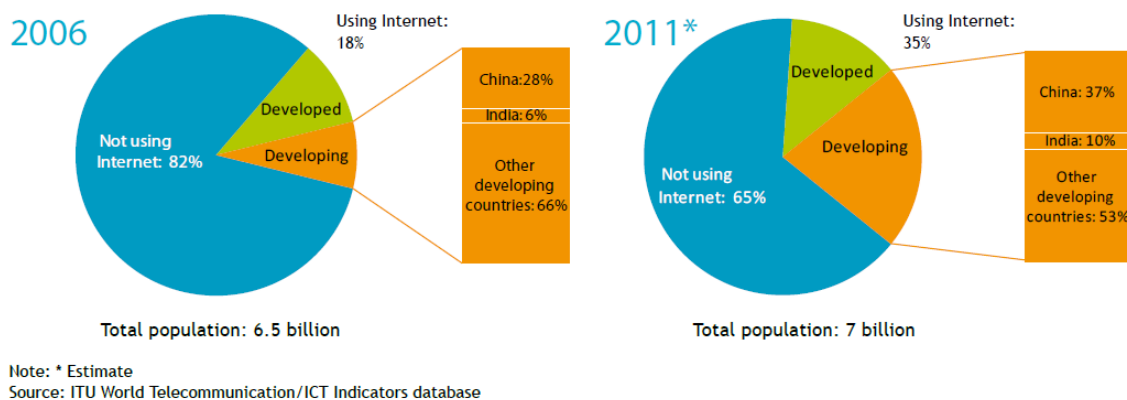


Ilustração 3. Crescimento da Internet nos Últimos Anos. Fonte: [ICT 2011].

Os serviços que a *internet* disponibiliza são: correio eletrónico, WWW (*World Wide Web* ou simplesmente *Web*), partilha de ficheiros, chamadas através da *internet*,

¹³ É uma ligação de documentos *web* textuais que facilita a navegação dos internautas.

¹⁴ É a fusão entre o hipertexto e a multimédia (imagens, sons e vídeo), que apresenta características de ambos e características próprias e únicas emergentes desta mesma fusão.

acesso remoto, entre outros. Entre todos, o serviço que mais se destaca é a WWW, uma vez que providencia informação hipermédia, estruturada em documentos HTML, chamadas páginas¹⁵. Uma coleção de páginas armazenadas na mesma diretoria de um servidor é conhecida como *site*. Através de uma qualquer página é virtualmente possível navegar até qualquer página existente em qualquer computador ligado à internet. Os documentos HTML são formados por texto estruturado com referências hipertexto e códigos de formatação. Toda esta informação é visualizada nos *browsers*, que convertem a estrutura HTML em instruções de visualização, otimizadas de acordo com a plataforma. Alguns dos formatos de dados que o HTML permite são: áudio (.wav, .mid, etc.), imagem (.gif, .jpg, .tif, etc.) e vídeo (.mpg, .avi, etc.). O protocolo usado para transferência de informação hipermédia é o HTTP (*Hypertext Transfer Protocol*), que corre por cima do TCP/IP (*Transmission Control Protocol/Internet Protocol*). Para aceder a este serviço só é necessário digitar o endereço URL (*Uniform Resource Locator*) no *browser* [Costa et al. 1995].

2.2. Motores de Busca

O crescimento da WWW tem sido tal que as tecnologias dos motores de busca aumentaram significativamente. Um dos indícios remonta-se ao ano 1994, quando um dos primeiros motores de busca indexou cerca de 110,000 páginas e documentos *web*, e outro em Novembro de 1997, quando os principais motores de busca afirmaram ter indexado entre 2 a 100 milhões de documentos *web* [Brin e Page 1998]. Os motores de busca em geral, realizam frequentemente uma verificação incremental da *web* para gerar vários índices relevantes, que podem ser pesquisados mais tarde, em resposta a uma consulta¹⁶ do utilizador.

Um estudo interessante sobre o crescimento da *web* é de [Gulli e Signorini 2005] que alega que o Google tem indexado mais de 8 bilhões de páginas, MSN¹⁷ cerca de 5 bilhões, Yahoo aproximadamente 4 bilhões e ASK um pouco mais de 2 bilhões no ano 2005. Dados como estes convertem o Google, Yahoo e Bing nos motores de busca mais

¹⁵ Uma página (conhecida comumente como página *web* ou *webpage* no Inglês) é um documento HTML com ligações de hipertexto que permitem navegar de uma página, ou secção para outra.

¹⁶ Consulta real introduzida por um utilizador na caixa de busca dos motores de busca.

¹⁷ Foi o nome inicial do motor de busca Bing.

populares de todos os tempos. Um recente estudo do *ComScore* nos EU pode corroborar esta afirmação: o Google é apresentado como líder dos motores de busca com uma quota de mercado de 66.8%, seguido pelo Bing com 15,6% e pelo Yahoo com 13,0% [ComScore 2012]. Outros dados estatísticos idênticos são apresentados pelo StatCounter, que atribui ao Google uma quota de mercado de 80%, ao Bing 9,5% e ao Yahoo 8% [StatCounter 2012]. Como se pode ver, o Google continua a ser o líder dos motores de busca, possuindo uma média de 70% de audiência nos Estados Unidos e 91% na Europa Ocidental [Lieb 2009].

Um motor de busca define-se como um fragmento de *software* que utiliza aplicações para recolher informações sobre os *sites*. As informações recolhidas são normalmente palavras-chave ou frases que são indicadores que se encontram no *site*, o URL, o código, os *links* internos e os *links* externos. Toda esta informação é indexada e armazenada numa base de dados [Ledford 2009]. Outro sentido para um motor de busca é a de Kent [2008] que o define “como um *site* manipulado por visitantes ou utilizadores para procurar documentos pela *internet* e a todo o sistema utilizado para rastrear, armazenar e indexar documentos *web*, e conduzir pesquisas na *internet*”. A vantagem de tudo este sistema é que se pode organizar e pesquisar a enorme quantidade de informação existente na *web* de uma maneira mais rápida e eficiente. Sullivan tem dividido o motor de busca em dois tipos: motores de busca baseados no *crawler*¹⁸ e às directorias geridas por pessoas, sendo que estes dois tipos de motores de busca armazenam *sites* de forma completamente diferentes [Sullivan 2003].

A listagem dos *sites* nos motores de busca, como o Google, Yahoo ou Bing, baseados no *crawler* é automática. Os motores de busca referidos rastreiam os *sites*, armazenando-os e indexando-os em bases de dados gigantescas, a fim de permitir que os utilizadores possam encontrar a informação de que estão à procura (Ilustração 4). Estes motores de busca são capazes de indexar uma variedade de tipos de conteúdo da *web*, tais como pdf, doc, ppt, xls, entre outros [Fister 2007], além de que podem encontrar com muita facilidade as alterações que os *sites* fazem, quer no seu código, quer na sua conceção. Porém, devido à enorme quantidade de páginas na *web* as indexações nem sempre estão dirigidas para o *site* todo [Kobayashi e Takeda 2000], são diversos os fatores que impedem os motores de busca realizar esta ação, como por

¹⁸ Conhecidos também como *spiders* ou robôs, são programas ou *scripts* automatizados que recolhem informação sobre cada URL da *web*. O *spider* do Google é o Googlebot, do Bing é o Bingbot e do Yahoo é o Yahoo Slurp 3.0.

exemplo, conteúdo duplicado, manipulação do código fonte, entre outros que não destacam a qualidade de um *site* [Lewandowski 2005]. No que respeita às alterações nos *sites*, de uma ou de outra forma, estas afetam a visibilidade nas SERPs, fazendo com que os *sites* fiquem mais longe do seu posicionamento inicial. O URL, as *Meta Tags*¹⁹, o conteúdo e outros elementos dos *sites* também desempenham um papel importante no posicionamento das páginas, no entanto, a maioria dos motores de busca utilizam uma mistura de *Meta Tags*, conteúdo, *Link Popularity*²⁰, *Click Popularity*²¹ e de longevidade para determinar o posicionamento dos *sites* [Curran 2004].

Na diretoria gerida por pessoas, comumente conhecido como “diretórios”, a listagem dos *sites* depende das pessoas sendo Dmoz²² um bom exemplo para abraçar esta abordagem. Neste tipo de motores de busca, os *webmasters* são responsáveis por enviar uma breve descrição dos seus *sites* para que estes possam ser armazenados nas suas bases de dados, ou os editores responsáveis por gerir estes diretórios escrevem uma pequena descrição para que estes sejam revistos depois. A procura dos utilizadores sobre estes motores de busca, corresponde apenas às descrições apresentadas e quando um *site* é alterado quer no seu conteúdo ou quer na sua conceição, não tem efeito sobre o posicionamento. As técnicas que são úteis para melhorar o posicionamento nos motores de busca baseados no *crawler* não funcionam nas diretorias geridas por pessoas. A única exceção é que um bom *site* com conteúdo único tem mais probabilidade de ser revisto do que um site com pouca informação [Sullivan 2003].

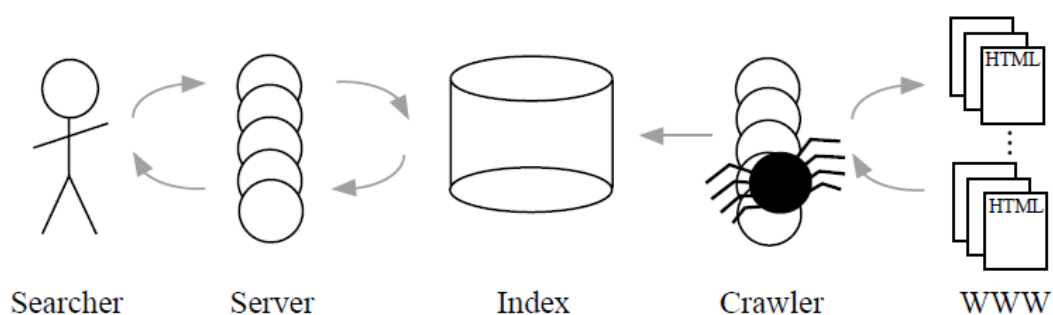


Ilustração 4. Motores de Busca Baseados em *Crawler*. Fonte: [Pinkerton, 2000].

Os motores de busca ainda têm estabelecido algumas diretrizes para ajudar os *webmasters* a maximizar o valor dos seus *sites* de uma maneira mais simples. O seu

¹⁹ São elementos HTML ou etiquetas que informam aos *crawlers* dos motores de busca sobre o *o title tag* (nome), *description tag* (descrição) e as *keywords tag* (palavras-chave) de uma página.

²⁰ Número e qualidade de *links* apontando a um *site*.

²¹ Mede o número de vezes que os utilizadores clicam num *site* através de motores de busca.

²² Deriva de *Directory Mozilla*, é um diretório de conteúdo aberto do *Nestcape*, o qual é gerido por editores voluntários.

propósito é que os *sites* possuam qualidade na sua conexão, código e conteúdo, apresentando informação útil e rápida aos utilizadores. No apêndice são apresentados as diretrizes dos maiores motores de busca: Google, Yahoo e Bing.

2.3. Search Engine Optimization (SEO)

SEO surge do acrónimo “Search Engine Optimization” ou “Search Engine Optimizer” [Google 2012]. Este último é adotado pela indústria de consultoria para se referir à pessoa responsável por realizar projetos de otimização dos *sites*.

Como já se tinha apresentado no Capítulo 1 são várias as definições que lhe atribuem, há quem o considere como um conjunto de técnicas, alguns como uma prática, outros como um processo, uns como uma arte e, ainda, outros como uma ciência. O termo Search Engine Optimization em si, não é mais que um conjunto de técnicas que se tem de aplicar em um *site* com a finalidade de o colocar o mais próximo possível do topo das páginas de resultados dos motores de busca para serem encontrados pelos utilizadores, ou seja, quando um utilizador digita uma frase ou uma palavra-chave na caixa de busca e recebe como resultado milhões de *sites*, o mais provável é que esteja a interagir com *sites* que utilizam as técnicas SEO.

O crescente uso da *internet*, a inumerável quantidade de informação na *web*, e o surgimento de novas tecnologias tem provocado uma mudança considerável no mundo dos negócios. É tal que, as empresas têm sido forçadas a interagir com os seus clientes de forma mais personalizada. É assim que estratégias como o *marketing* se tem adaptado à era digital, tentando promover os seus produtos e serviços, atendendo às necessidades dos seus utilizadores e desenvolvendo novos e potenciais clientes através da *web*, daí surge o nome de *Marketing Online*. Para [Gay et al. 2007] o termo refere-se ao processo de construção e manutenção do relacionamento com o cliente por meio de atividades *online* para facilitar o intercâmbio de ideias, produtos e serviços que satisfaçam os objetivos de ambas partes. Atualmente é constituído por diferentes subáreas como: SMM (Social Media Marketing) [Tuten 2008], SEM (Search Engine Marketing) [McDonald e Zayat 2009], *E-mail Marketing*, *Affiliate Marketing* [Spindler 2010], *Referral Marketing* [Thurow 2007], SMO (Social Media Optimization)

[Miletsky 2009], e SEO [Jones 2008], sendo SEO o componente mais relevante de *Marketing Online*.

As páginas de resultados de motor de busca (SERPs) normalmente contêm uma mistura de resultados de busca naturais ou orgânicas e resultados de busca paga (Ilustração 5). Geralmente, mas nem sempre, os resultados de busca paga distinguem-se dos resultados de busca naturais de forma que o visitante pode reconhecer quais listagens são resultados naturais e quais resultados são pagos [Work 2008]. Google, Yahoo e Bing empregam um modelo pago CPC (Cost Per Click) onde os *marketers* ou anunciantes têm a possibilidade de transacionar sobre a quantia que estão dispostos a pagar por cada clique que o seu resultado patrocinado recebe de cada palavra-chave que desejam atingir. Este modelo pago gera receitas aos motores de busca, parte das quais permitem desenvolver a busca orgânica. Por outro lado, a busca orgânica não utiliza este modelo pago para determinar que resultados serão apresentados ao visitante, geralmente, os resultados de busca orgânica são exibidos usando um algoritmo baseado na qualidade de um *site* e não podem ser influenciados por dinheiro [McVittie 2012].

PPC (Pay per click) tem diferentes nomes que significam a mesma coisa: *Pay per placement*, *Pay per ranking*, *Pay per position* [Arnold 2010], CPC, busca paga, busca patrocinada e SEM. Este último também é utilizado para se referir ao *Marketing* baseado no motor de busca em um sentido mais geral, que também abrange SEO. No entanto, em um sentido mais restrito, SEM refere-se ao *Marketing* Pay per click [Work 2008] e SEO refere-se ao *Marketing* de busca orgânica, normalmente conhecido como Organic SEO [McDonald e Zayat 2009].

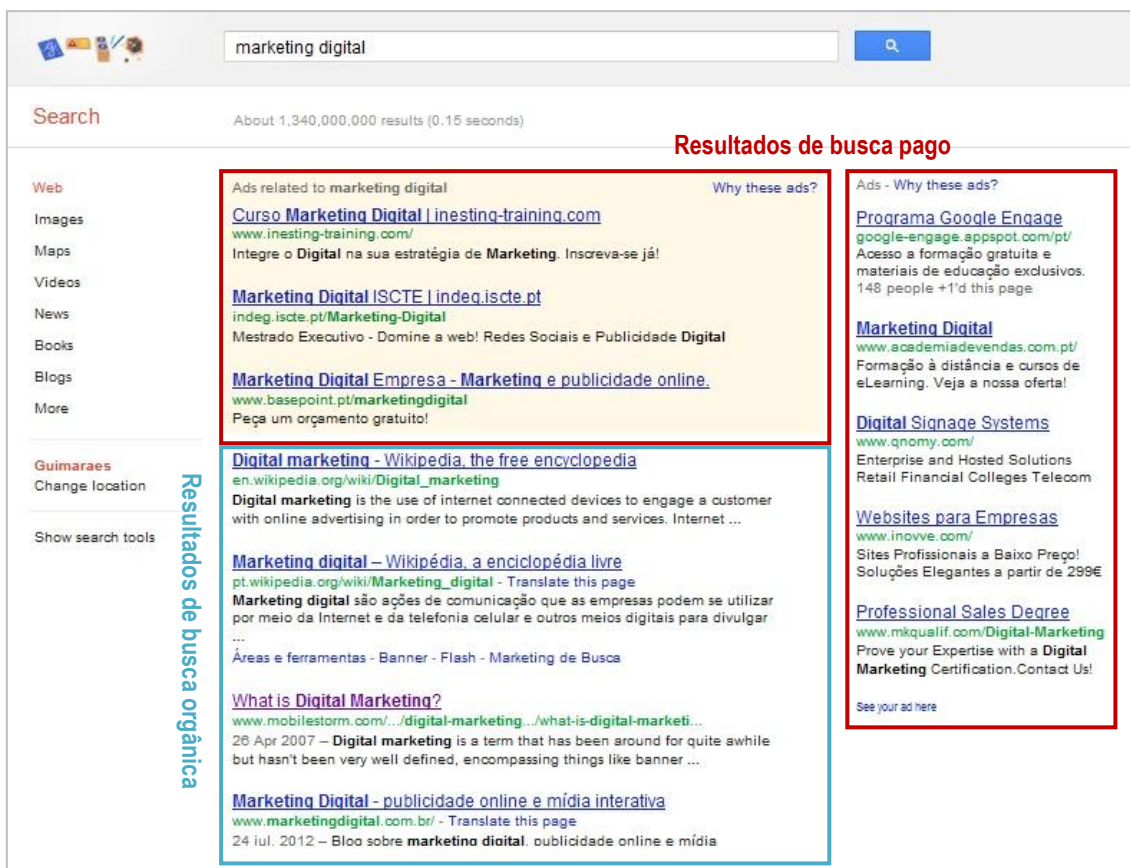


Ilustração 5. Resultados Orgânicos e Pagos nos Motores de Busca.

Processo SEO

Segundo as definições dos parágrafos anteriores, SEO é um conjunto de técnicas que visam gerar Tráfego para posicionar um *site* nas SERPs, sendo que estas técnicas incluem a manipulação de dezenas ou mesmo de centenas de elementos do site. O processo SEO está dividido em quatro fases, no entanto, como o SEO está a sofrer mudanças contínuas, as fases do processo podem mudar ao longo do tempo. O processo habitual é: Pesquisa e seleção de palavras-chave/frases, Otimização *On-Page*, Otimização *Off-Page* [Malaga 2008] e Monitorização e Apresentação de Relatórios [McVittie 2012] (ver Ilustração 6).

Fase 1. Pesquisa e seleção de palavras-chave/frases

Durante esta primeira fase desenvolve-se uma lista de palavras-chave e/ou frases específicas com o objetivo de criar referências relevantes para os motores de busca.

Estes termos fazem com que um *site* apareça nas SERPs quando o utilizador os digita nos motores de busca. Além de desenvolver uma lista de palavras-chave e frases, também se determina a competitividade de cada termo e a frequência com que é utilizado numa busca [Malaga 2008]. A pesquisa de palavras-chave não se aplica apenas aos elementos textuais *on-page*, também se aplica para escolha do nome de domínio, *links* externos, composição de *links*, listagens nas diretorias, e para muitos outros elementos [Jerkovic 2009].

Fase 2. Otimização On-Page

Durante esta segunda fase, cada página do *site* é otimizada pelo *webmaster* através da manipulação de vários elementos *on-page* [Malaga 2008]. Os elementos *on-page* são todas as técnicas de otimização que podem ser realizadas diretamente no próprio *site*, incluído a sua idade [Jerkovic 2009], e que não podem ser influenciadas por fatores externos [Spindler 2010]. Estes elementos *on-page* [David 2011] são:

- Conteúdo do corpo de texto principal da página.
- Título e *Meta Tags*.
- *Headings Tags* (H1, H2, H3).
- Qualidade e complexidade do código HTML e CSS (Cascading Style Sheets) que gera o *site*.
- Imagens, nomes dos ficheiros e *Alt Tags*.
- Atributos de texto como o uso em negrito e sublinhado.
- *Links* internos, com o seu respetivo número e texto âncora²³.
- Utilização dos atributos *dofollow* ou *nofollow* em qualquer dos *links*.
- Menus, navegação interna e estrutura de *links*.
- Tamanho de ficheiros e velocidade com que o *site* carrega.
- Número total de páginas no *site*.
- Índice de atualização de conteúdo do *site*.

²³ Texto visível dentro de um *link*.

Fase 3. Otimização *Off-Page*

Nesta terceira fase, os maiores motores de busca consideram o número e relevância de *links* de sites externos que referenciam um *site*, ou seja, são considerados os elementos *off-page*. Os elementos *off-page* são ações que não podem ser executadas no próprio *site* mas sim em outros sites [Spindler 2010]. Portanto, os projetos de SEO geralmente incluem uma fase de construção de links chamado *Link Building*²⁴. É aqui onde os *webmasters* ou responsáveis de sites solicitam *links* a outros *webmasters* e muitos deles usam aplicações de *Link Building* [Malaga 2008]. Os elementos *off-page* podem ser resumidos numa palavra: *backlinks*. Os *backlinks* são a verdadeira força que faz classificar os *sites*. Em mercados competitivos de busca, os *links* abarcam cerca de 80 a 90 % do trabalho que se faz num *site*. *Sites* com maior número de *backlinks* são mais propensos a serem posicionados nos primeiros lugares das SERP do que aqueles que têm um menor número de *backlinks* [David 2011]. Os motores de busca são responsáveis por fazer medições deste tipo de elementos, por exemplo o Google usa o *PageRank*²⁵ para analisar os *backlinks* de um *site* [Ascensão 2011].

Fase 4. Monitorização e Apresentação de Relatórios

Nesta fase do processo, procura-se perceber a reação dos motores de busca perante o trabalho SEO. A monitorização pode incluir a revisão da classificação do *site* destino nos motores de busca, revendo os *links* que apontam para o *site*, e monitorando o tráfego *web* e a análise de conversão²⁶. SEO é um processo iterativo contínuo sem um ponto inicial e final diferente. Os algoritmos dos motores de busca e *sites* concorrentes não permanecem estáticos, portanto, um *site* deve ser monitorado numa base contínua para garantir que o *site* conserva uma visibilidade de busca alta. Os resultados da monitorização contínua e a apresentação de relatórios são usados para influenciar a próxima iteração de pesquisa de palavra-chave e as outras fases do processo [McVittie 2012].

²⁴ Uma das principais técnicas *White Hat* SEO.

²⁵ É o algoritmo registado e patenteado pelo Google que mede a relevância dos *sites* na *web* através da análise de *links* externos.

²⁶ É a capacidade que um *site* tem de manter os utilizadores, ou seja, convertê-los de visitantes em clientes.

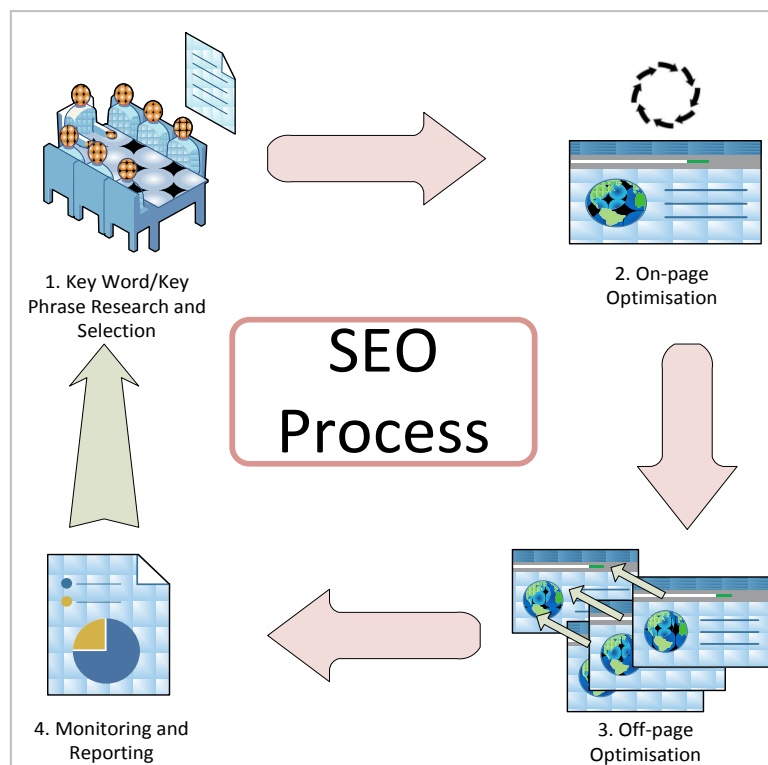


Ilustração 6. Processo SEO. Adaptado de Malaga [2008] e McVittie [2012].

2.4. Técnicas SEO

O SEO possui um conjunto de técnicas que são empregadas no momento de otimizar um *site* para obter visibilidade nas páginas de resultados dos motores de busca. Algumas delas estão aprovadas pelos motores de busca enquanto outras tentam burlar o sistema. Vários autores dividiram estas técnicas em dois grupos: Técnicas *White Hat* SEO e Técnicas *Black Hat* SEO [Bröcker e van Ahee 2008; McDonald e Zayat 2009; Svore et al. 2007]

2.4.1. Técnicas *White Hat* SEO

As técnicas *White Hat* SEO foram criadas a pensar nos utilizadores finais, mas de forma que os indexadores dos motores de busca possam facilmente navegar no *site*. Estas técnicas, também chamadas técnicas éticas, seguem as diretrizes sugeridas pelos motores de busca [John et al. 2011] e focam-se em melhorar a qualidade e conteúdo de uma página de modo que a informação sobre a página seja útil para muitos utilizadores [Svore et al. 2007]. Quando as técnicas *White Hat* SEO são violadas em um *site* tornam-

se rapidamente desaprovadas pelos motores de busca [McDonald e Zayat 2009]. A seguir apresentam-se as técnicas *White Hat* SEO mais relevantes, sendo que cada técnica tem sido agrupada numa de três categorias: conteúdo, código e ligações.

1. Conteúdo

1.1. Conteúdo Próprio

Produzir um bom conteúdo que seja original e bem escrito é absolutamente fundamental para o sucesso a longo prazo de SEO. Nesta nova era da Internet um bom conteúdo é viral²⁷ [Fleiss 2007], é fundamental para que os visitantes determinem o valor de um *site*, e é um dos principais fatores que os motores de busca usam para a classificação. Apesar de se dever manter os princípios de SEO ao criar conteúdo em mente, a chave para construir uma classificação de longo prazo nos motores de busca é, escrever conteúdo para as pessoas e não para os motores de busca [Jones 2008].

No entanto, o conteúdo muitas vezes é ignorado pelos *webmasters*, sendo que geralmente eles colocam pouco ou nenhum texto, algumas imagens ou *sites* com *flash*. Os motores de busca não lêem texto nas imagens ou em *flash*. De modo que, se um serviço está a ser descrito em ficheiros de imagem ou *flash* a mensagem não será lida e não será classificada para esses termos [David 2011]. O texto sobre a página deve cumprir os critérios seguintes:

- **Tamanho:** Uma página deve ter um mínimo de 250 palavras legíveis de texto ASCII.
- **Enfoque:** Uma página deve estar focada num conjunto de palavras-chave razoavelmente limitado.
- **Densidade de Palavras-chave:** Uma página não deve ter palavras-chave repetidas de modo que a densidade das palavras-chave sejam muito altas em relação ao número total de palavras.

²⁷ Tornar-se muito popular, circulando rapidamente de pessoa para pessoa, especialmente através da *web*. Praticamente como um vírus.

1.2. Comunidades

Criar uma comunidade, como um *blog* ou fórum num *site* é uma das técnicas mais eficazes de manter o conteúdo atualizado, ajudando a estabelecer o *site* como uma autoridade para uma determinada área de negócio. Os motores de busca favorecem *sites* com conteúdo atualizado mais do que *sites* que não atualizam o conteúdo tão frequentemente. Iniciar uma comunidade é fácil e barato, o benefício disso é promover a interação e criação de conteúdo entre os utilizadores [Jones 2008].

Do ponto de vista de conteúdo, a Internet tornou-se muito mais fácil de navegar seguindo o princípio de RSS (*Really Simple Syndication*). O RSS permite que um programa chamado agregador (ou leitor de *feeds*) notifique os utilizadores sobre novos conteúdos adicionados ao *site*, recupere esse conteúdo novo, e o exponha ao utilizador numa interface simples de usar. Os RSS e os *blogs* andam a par por causa da natureza constante de atualização dos *blogs* [Fleiss 2007].

1.3. RSS Feeds

RSS (*Real Simple Syndication or Rich Site Summary*) está a tornar-se numa ferramenta poderosa para os *marketers* da Internet. Os visitantes podem adicionar facilmente e rapidamente novos conteúdos ao *site*. Os *feeds* de artigos são atualizados com frequência, de modo que se possa providenciar aos visitantes (e aos motores de busca) o que eles querem: “conteúdo recente”. A finalidade desta técnica é promover qualquer conteúdo novo, como novos produtos, ofertas especiais, artigos ou opiniões [Enaceanu 2008].

2. Código

2.1. Tagging

O conteúdo é um elemento essencial para o *design* do *site*, mas é uma parte do *design* dirigida ao cliente, enquanto o HTML é uma parte dirigida ao rastreador. Antes que os clientes vejam o conteúdo, os *crawlers* irão ver o HTML. O *ranking* SEO dependerá, em grande parte dos *tags* que controlam a página. Os clientes nunca vêem os *tags*, mas sem eles, o *site* não será visto [Ledford 2009].

- **Title Tag:** Trata-se do *tag* mais importante numa página [Princz 2007]. No *title tag* deve colocar-se apenas as palavras-chave exatas que estão a ser utilizadas para otimizar o *site*. Cada página deve ter o seu próprio *title tag* [Enaceanu 2008].
- **Meta Tags:** Os *Meta Keywords*, *Meta Description* e o *ALT* foram muito importantes há alguns anos. Os *webmasters* tentaram encher esses *tags* com muitas palavras. Hoje em dia, por causa de problemas de *spam*, os *meta tags* nem sequer são lidos por alguns motores de busca [Princz 2007].
- **Tags de conteúdo:** *Tags* como cabeçalho <H1> <H2> <H3>, negrito , itálico <I>, sublinhado <U>, fortemente enfatizado , entre outros, são importantes [Princz 2007].

2.2. Linguagem de Estilo

Uma técnica amplamente utilizado no *design* de *sites* que permite separar a formatação de uma página do seu conteúdo é a linguagem de estilo CSS. O CSS faz que as *tags* de formatação sejam estritamente eficazes nele enquanto as outras *tags* fazem realmente o trabalho necessário de obter o *site* listado naturalmente no motor de busca [Ledford 2009]. Um dos seus principais benefícios é a facilidade de atualização das páginas, uma vez que só é preciso atualizar o ficheiro CSS em vez de cada página do *site*. O seu uso nos *sites* faz do código fonte muito mais ágil permitindo aos *crawlers* dos motores de busca

rastreá-los de forma mais simples. A técnica aplica-se geralmente a páginas escritas em HTML, XHTML ou qualquer tipo de documento XML.

2.3. URLs Curtas

Idealmente devem manter-se as URLs curtas com nomes de página que tenham as palavras-chave que reflitam melhor o conteúdo da página. Um delimitador como sublinhado ou traço para separar as palavras-chave ou modelo dos produtos uns dos outros é uma boa opção [Enaceanu 2008].

2.4. Criação de *Sitemap*²⁸ XML

Uma das técnicas para conseguir uma rápida indexação nos motores de busca é criar um mapa do *site*. Mas este não é o mapa do *site* comum que ajuda os utilizadores a navegarem rapidamente através de uma página *web*, mas sim é um documento baseado em XML que contém informações como o URL, a última atualização do *site*, a relevância de páginas vizinhas, entre outros, sobre cada uma das páginas que contem um *website*. Por outro lado, um mapa do *site* não garante que todas as páginas sejam rastreadas ou incluídas nas SERP, porém pode dar alguma direção aos *crawlers* de modo a melhorarem o rastreamento nos *sites* [Ledford 2009].

Apos de ter sido criado o mapa do *site* XML, o documento é inserido no ficheiro robots.txt para os *crawlers* o encontrem ou para serem submetidos ao motor de busca com que se pretende trabalhar. O ficheiro de mapa do *site* não pode conter mais de 50.000 URLs e não deve ser maior do que 50MB quando é descomprimido [Google Webmasters, 2012]. Uma das formas de submeter este ficheiro nos motores de busca é ter uma conta neles, por exemplo, uma conta no Google implica ter acesso a vários dos seus serviços, um deles é a ferramenta Google *Webmaster Tools*²⁹ que permite submeter e verificar o ficheiro *sitemap* XML.

²⁸ É um ficheiro XML que lista os URLs de um *site*.

²⁹ É uma ferramenta do Google que permite aos *webmasters* verificarem o estado de indexação e otimizar a visibilidade dos sites.

3. Ligações

3.1. *LinkBuilding*

A construção de uma grande estrutura de *links* para um *site* não é algo que acontece no tempo, é necessário que isto ande a par com o *site* desde o seu lançamento. Uma construção de *links* bem-sucedida pode levar meses, e às vezes mais. Algo que se tem de considerar neste processo é que a quantidade não é tão importante quanto a qualidade [Ledford 2009]. A todo este processo de construção de *links* chama-se *linkbuilding*.

A. *Links Internos*

Trata-se de *links* que vão desde uma página *web* para outra página diferente dentro do mesmo domínio [SeoMoz 2012], note-se que é diferente da estrutura de navegação. Os *links* internos são os *links* mais naturais que ocorrem no texto de um *site*. Sem uma estratégia de *links* internos, o *site* não será rastreado corretamente pelos motores de busca. Portanto sugere-se obter *links* que vão desde um elemento (como um *post* no *blog*) para outros elementos importantes (como um artigo arquivado ou um *clipping* de notícias no *site*), de modo que os visitantes se desloquem sem nenhum problema através do *site*. Os métodos mais eficazes de *links* internos são os *links* de texto, os *links* dentro dos rodapés de páginas e os *links* de texto *inline*. Os *links* de texto são aqueles que utilizam o texto âncora ou palavras-chave para aparecerem com frequência no texto de uma página. Os *links* dentro dos rodapés de páginas são os mais usados atualmente. Os *links* de texto *inline* são *links* contidos frequentemente no corpo de algum elemento do conteúdo de um *site*, tais como artigos e blogs [Ledford 2009]. Algumas das suas vantagens são [Enaceanu 2008]:

- Garantir que o *site* obtém o rastreamento necessário e que todas as páginas sejam encontradas pelos motores de busca.
- Fazer que uma página *web* consiga relevância através de uma palavra-chave específica.
- Aumentar o *PageRank* de uma página.

B. Links Externos

Os *links* externos subdividem-se em *Outbound Links* e os *Inbound Links* [Ascensão 2011].

1. Outbound Links

Este tipo de *links* são os mais simples de conseguir, uma vez que apontam desde o *site* a ser otimizado para outro *site*. O responsável do *site* tem o controlo complete e total sobre estes *links*, o seu principal objetivo é fazer que o *site* se beneficie com o tráfego mais elevado possível. Os *outbound links* desempenham um papel fundamental na estratégia geral de ligações e agregam valor aos *sites*, fornecendo informações úteis aos utilizadores sem necessidade de criar o conteúdo. No *blog* de SeoMOZ, Eric Enge apresenta um *post* onde sugere aos especialistas registar *sites* que ajudem aos próprios *sites*, especialmente aqueles que não são as da sua concorrência mais directa, e incita citá-los como fontes ao longo do conteúdo [SeoMoz Blog, 2008]. A finalidade deste trabalho é que os visitantes dêem maior valor ao *site* e ao mesmo tempo seja atraem-te para outros *sites* a fim de obter ligações externas. Enaceanu [2008] sugere que as páginas onde se *linkam* os outros *sites* não estejam situados para além de três cliques de distância desde a página inicial. Ainda menciona que o *site* a ser otimizado deve ligar *sites* com conteúdo de qualidade e tópico semelhante ou relacionado, a fim de contribuir à construção de um melhor tópico para o *site*. Também sublinha que cada página do *site* deve enlaçar à página inicial e aos principais serviços.

2. Inbound Links (Backlinks)

Também conhecido como *incoming links*, *inlinks*, ou *inward links*, são ligações recebidas para um *site* ou página da *web*. Estes *links*, antes do surgimento dos motores de busca, foram muito importantes, como principal meio de navegação na *web*, hoje, a sua importância reside nas técnicas SEO. O número de *backlinks* é uma indicação da popularidade ou da importância de um *site* ou página, por exemplo o Google usa este tipo de *links* para determinar o *PageRank* de um *site*. Além de SEO, os *backlinks* de um *site* podem ser de

interesse pessoal, cultural ou semântico, uma vez que eles indicam quem está a prestar atenção a essa página [Muljadi 2011].

3.2. Submissão a Directorias

Uma forma de conseguir uma rápida indexação nos motores de busca é submetendo os *sites* às directorias de busca [Malaga 2008]. As directorias de busca não são motores de busca nem se comportam como eles, mas recolhem e categorizam *links* para outros *sites*. Pode-se dizer que são uma forma fácil de obter *links* externos. Atualmente, existem directorias gratuitas e outras que são pagas. Uma das directorias mais conhecidas na *Internet* é o Open Directory Project, que é gerido pela dmoz.org [Ledford 2009].

3.3. LinkBaiting

Segundo Fleiss [2007] a técnica *linkbaiting* é uma das chaves para gerar bom conteúdo nos *sites*. O seu objetivo é chamar a atenção das pessoas sobre um *site* para este ser *linkado* por outros *sites*. A partir de um artigo polémico e bem escrito, até a um vídeo de um buldogue sobre um *skate*, os proprietários de *sites* ligarão todo e qualquer conteúdo enquanto ele for interessante e chame a atenção das pessoas. Não há limites para os tipos de conteúdos que alguém pode usar para atrair visitantes a um *site*. De fato, o próprio nome de um novo tipo de *linkbaiting* sugere uma qualidade difícil de definir. Esta nova tática de *linkbaiting* chama-se “*Widget Baiting*”. Esta diz-nos que um *widget* é um elemento de interface do utilizador que exhibe ou fornece informação de uma maneira específica para um utilizador interagir com uma aplicação.

Um *widget* pode ser uma agenda, um *ticker* de cotações, uma frase do dia, ou um ícone que coleciona os vídeos mais populares de YouTube. Na sua forma mais básica, um *widget* é uma ferramenta interativa, transferível, feita por peças simples de código que podem ser facilmente adicionados a um *site*. Quando um *widget* é adicionado a um *site*, caso seja codificado corretamente, vai agir como um *link* indexável que aponta de volta para a sua página de origem. Esses *links* podem ajudar a impulsionar um *site* nos motores de busca, mas também representam um grande potencial para o tráfego orgânico [Fleiss 2007].

2.4.2. Técnicas *Black Hat* SEO

Qualquer otimização que provoque que um *site* se posicione mais acima do que o seu conteúdo justifica, ou qualquer alteração feita especificamente para os motores de busca que não melhore a experiência do utilizador é considerada como técnica *Black Hat* SEO [George 2005] ou *Spamdexing* SEO [Schimmel, 2010]. Na realidade, a definição destas técnicas é um assunto de debate entre os peritos em SEO. Os praticantes destas técnicas afirmam que a ética não entra em SEO, que qualquer otimização efetuada para um *site* é feita como um esforço deliberado para manipular a página de resultados dos motores de busca [George 2005]. Por exemplo, do ponto de vista dos *sites* comerciais, estas técnicas são utilizadas para ganhar dinheiro pelo que tentam ser classificadas em lugares mais altos do que seus concorrentes dentro dos motores de busca. Com tudo isto, compete aos motores de busca definir parâmetros de otimização razoável.

A maioria das técnicas *Black Hat* SEO estão destinadas a fornecer determinado conteúdo só para os *crawlers*, enquanto os utilizadores vêem um conteúdo completamente diferente. A razão para que isto acontecer é o facto de ter o conteúdo que é utilizado para alcançar elevadas classificações pode não ser favorável para o *design* do *site* ou para a elevada taxa de conversão³⁰ [Malaga 2008]. A seguir apresentam-se as técnicas *Black Hat* SEO mais correntes. Estas técnicas estão agrupadas em três categorias, tal como as anteriores: conteúdo, código e ligações:

1. Conteúdo

1.1. Conteúdo Oculto

Inicialmente, os praticantes desta técnica usavam elementos HTML como primeiro plano e cores de fundo, ou fontes muito pequenas para adicionar conteúdo oculto aos *sites*. Recentemente alguns praticantes têm começado a utilizar folhas de estilo em cascata (CSS) para ocultar elementos. Os elementos escondidos são colocados dentro dos *Tags Div* ocultos. Porém, os motores de busca encontraram rapidamente estas técnicas e começaram a penalizar *sites* que as usavam. O Google, por exemplo, começou a remover conteúdos que

³⁰ Taxa dos visitantes que executam uma ação de monetização, tais como fazer uma compra. Uma das formas mais populares de fazer monetização num *site* é usando o Google *AdSense* que permite aos *sites* anunciar os seus produtos ou serviços a fim de obter lucros.

continham *Tags Div* ocultos do seu índice. A penalização e/ou a remoção de um *site* são soluções possíveis para os motores de busca lidarem com esta técnica, no entanto é um problema para alguns *webmasters* que utilizam *sites* com divisões ocultas nos seus projetos [Malaga 2008].

1.2. *Scraping Content*

Através desta técnica, o conteúdo de um *site* é plagiado e colocado em um domínio diferente [Marie 2012], isto é, quando um *site* copia o seu conteúdo a partir de outros sites está a incorrer nesta técnica. Os motores de busca não aprovam *sites* que plagam porque eles não são originais e porque geralmente direcionam aos visitantes para um *site* que é menos relevante para os seus termos de busca [Ledford 2009].

1.3. *Article Spinning*

É usada por *bloggers* para manipular a sua classificação. Artigos existentes são *postados* novamente usando elementos ou conteúdo ligeiramente diferente do original, a fim de ser indexado novamente como conteúdo novo, em vez de receber penalidades por *postar* conteúdo duplicado [Schimmel, 2010].

2. Código

2.1. *Doorway Pages*

O objetivo desta técnica é atingir altas classificações nos motores de busca para várias palavras-chave ou frases. O praticante desta técnica cria uma página separada para cada palavra-chave ou frase. Alguns praticantes utilizam centenas destas páginas. As *Doorways Pages* costumam utilizar um *Meta Refresh* rápido para redirecionar os utilizadores à página principal. O *Meta refresh* é um comando HTML que alterna automaticamente os utilizadores para outras páginas depois de um período de tempo especificado, e é utilizado em *sites* fora da data. Atualmente os principais motores de busca removem páginas que contenham *Meta Refresh* [Malaga 2008].

2.2. Cloaking

É uma técnica frequentemente utilizada para tentar enganar os motores de busca e os utilizadores. Isto é, a versão de um *site* que o utilizador está a ver é diferente da que os motores de busca vêem quando eles estão a rastrear o *site*. Uma das razões mais comuns que levam os praticantes a utilizarem esta técnica é porque o *site* ou página específica está composta por componentes que os motores de busca não podem ler. Neste sentido, o *site* tenta posicionar-se de forma lícita, mas revela outro *site* quando o utilizador entra nele [Marie 2012], por exemplo um *site* pornográfico, mostrando assim duas caras.

2.3. Keywords Stuffing

Esta técnica está focada em encher palavras-chave no rodapé ou no conteúdo de uma página. No entanto, os motores de busca não classificam *sites* deste tipo, pois são baseados unicamente na frequência de palavras-chave, e se eles vêem muitas palavras-chave numa página podem penalizá-lo (incluindo uma queda na posição) ou excluí-lo da lista dos SERP em casos graves [Marie 2012].

2.4. Páginas PPC

O objetivo desta técnica é criar páginas exclusivamente para colocar anúncios PPC do Google AdSense³¹, com pouco ou nenhum conteúdo útil [Kent 2008].

3. Ligações

3.1. Links enganosos

É uma técnica onde o domínio de um *site* aparentemente conduz a um lugar, mas, na verdade, conduz a outro. Um exemplo de um *link* enganoso é aquele que se mostra como www.onewebsite.com, mas na verdade está a conduzir a iquestionmark.com [Ledford 2009].

³¹ Um serviço de publicidade oferecido pelo Google.

3.2. *Links Farm*

Esta técnica é usada para criar muitos *sites* com pouco ou conteúdo semelhante, os quais apontam para um *site* de origem que está a ser promovido [Metaxas e DeStefano 2005]. Os *sites* que utilizam esta técnica podem enganar os sistemas de busca de informação, fazendo-os acreditar que o *site* de origem é uma boa página web, uma vez que vários *links* apontam para ele. [Langville e Meyer 2003]. Os *links farm* geralmente são criados utilizando ferramentas automáticas [Malaga 2008].

3.3. *Wiki Spam*

Esta técnica usa a capacidade de edição aberta dos *sites Wiki* para adicionar *links* a um *site spam* na lista dos *Wiki*. Esta prática tornou-se muito conhecida e foi utilizada constantemente pelos *Spammers*, especialmente na *Wikipédia*, tendo sido por isso adicionado o *tag Nofollow* na maioria das novas introduções *Wiki* para impedir *Spam* nos motores de busca [Ledford 2009].

3.4. *Blog/Forum Spam*

Assim como nos *wikis*, os *links spam* em comentários de *blogs* e fóruns são bem divulgados. Estima-se que 75% dos comentários nos *blogs* sejam *spam* [West et al. 2012]. Quando os *blogs* se tornaram populares, muitos *webmasters*, começaram a visitá-los e a “postar” mensagens com *links* que apontavam para os seus *sites* a fim de promovê-los e posicioná-los nos motores de busca. Por isso hoje, na maioria dos *blogs* e fóruns utiliza-se o *tag Nofollow* para pôr fim ao *Spam*.

2.5. Ferramentas de *Web Analytics*

Um *site* pode ser usado para obter informações estatísticas sobre os visitantes e permitir às empresas rastrear e segmentar os seus clientes, bem como acompanhar os seus comportamentos de compra. Uma boa forma de fazer isto é usando ferramentas de *web analytics* que proporcionam informação valiosa sobre o tráfego do *site* e a eficácia do *marketing* [Google Analytics 2012].

Web Analytics, também conhecida como *Web Metrics*, *Web Measurement* ou *E-Metrics* [Pineda et al., 2004], define-se como a medição, recolha, análise e comunicação de dados da *internet* com a finalidade de compreender e otimizar o uso da *web* [DAA 2008], os seus objetivos primários são: melhorar o desempenho do *site* tanto do ponto de vista técnico como de conteúdo, melhorar a experiência do visitante, contribuir para a compreensão global de clientes e canais, e identificar oportunidades e riscos [Peterson 2004].

A *Web Analytics* tem uma enorme quantidade de métricas, entre elas encontram-se as métricas tradicionais e as métricas de valor agregado. As métricas tradicionais concentram-se apenas na medição de tráfego e as métricas como *hits*, páginas vistas ou número de visitas possuem um valor agregado para o negócio [Pineda et al. 2004]. No entanto, durante o segundo semestre de 2006, o comité de normas da Digital Analytics Association iniciou um esforço para definir aquilo que foi acordado como as três métricas mais importantes na *web*: visitantes únicos, visitas/sessões e páginas vistas [DAA 2008].

Assim, ferramentas de *Web Analytics* são essenciais para melhorar os *sites* [Enaceanu 2008]. No mercado atual pode-se encontrar uma grande variedade dessas ferramentas, entre gratuitas e pagas. A seguir apresenta-se uma listagem delas com algumas das suas vantagens e desvantagens (Tabela 1):

Ferramentas da <i>Web Analytics</i>	Vantagens	Desvantagens
Webtrends http://www.webtrends.com Preço: \$10.000+	<ul style="list-style-type: none"> ✓ Pacote analítico poderoso ✓ Oferece tudo o que oferecem outras ferramentas, e muito mais 	<ul style="list-style-type: none"> ✓ Caro ✓ Difícil para utilizadores inexperientes
Analog http://www.analog.cx Preço: Gratuito	<ul style="list-style-type: none"> ✓ Sólida solução ✓ Popular 	<ul style="list-style-type: none"> ✓ Tudo é mostrado em um único ficheiro HTML
WebLog Expert http://www.weblogexpert.com Preço: \$74.95-\$124.95	<ul style="list-style-type: none"> ✓ Interface elegante com gráficos ✓ Diferentes estilos de relatório (.html, Pdf, etc) ✓ Relativamente barato 	<ul style="list-style-type: none"> ✓ Não tem características de ferramentas de gama alta
AWStats http://awstats.sourceforge.net Preço: Gratuito	<ul style="list-style-type: none"> ✓ Interface de aparência profissional ✓ Popular ✓ Codificado em Perl 	<ul style="list-style-type: none"> ✓ Pode ter alguns gráficos adicionais encontrados em ferramentas pagas
Webalizer http://www.webalizer.com Preço: Gratuito	<ul style="list-style-type: none"> ✓ Rápido, codificado em C ✓ Interface simples ✓ Abordagem direta 	<ul style="list-style-type: none"> ✓ Leva algum tempo para se acostumar com diagramas de métricas empilhadas ✓ Pode ter alguns gráficos adicionais em AWStats

Tabela 1. Ferramentas Estatísticas mais Populares. Adaptado de [Jerkovic 2009].

No entanto, os próprios motores de busca também oferecem ferramentas de *Web Analytics*, por exemplo, a Google oferece uma grande ferramenta de análise gratuita chamada *Google Analytics*, a Yahoo outra ferramenta similar chamada *Yahoo! Web Analytics*, e a Microsoft com algumas características similares às ferramentas de *Web Analytics*, oferece o *Bing Webmaster Tool*.

Google Analytics

Quando o Google lançou o *Google Analytics* em 2005 como ferramenta gratuita revolucionou a *web analytics*. A análise da *web* já não era mais um projecto caro que exigia um investimento substancial em *software*. Naquele momento, as ferramentas de medição tornaram-se irrelevantes do ponto de vista de custo, e as pessoas começaram a olhá-las de uma maneira diferente. Com a criação de uma ferramenta de *web analytics* gratuita que é fácil de entender, o Google potenciou que todos se tornassem em analistas

web, e hoje, *web designers*, *webmasters*, equipas de TI, executivos de nível C, e *marketers*, todos usam o *Google Analytics* para monitorar e medir o desempenho dos seus *sites* e as iniciativas de *Marketing Online* [Cutroni 2010].

Google Analytics é uma ferramenta poderosa para medir o sucesso de um *site*, os esforços de *marketing*, os produtos e os serviços em que o utilizador está a usufruir [Ledford et al. 2010]. Alguns outros autores definem-no como uma solução de análise gratuita projetada para dar uma visão completa do desempenho de um *site*. Compreender a maneira correta de analisar e implementar os vários tipos de dados que o *Google Analytics* permite aos *webmasters* ter uma vantagem considerável sobre os concorrentes na missão de classificar os *sites* no topo da busca orgânica. Além de que o *Google Analytics* fornece uma visão de alto nível do tráfego no *site* e da interação com o utilizador, o que permite analisar as diversas fontes de tráfego que chegam ao *site*. O *Google Analytics* é gratuito e fácil de instalar. O detalhe mais importante no processo de instalação é ter certeza de que o código de rastreamento seja colocado corretamente em cada página do *site*. Isto garante que todos os relatórios do *Google Analytics* sejam os mais precisos e confiáveis possíveis [Jones 2008].

As métricas que se podem extrair de um *site* dependem dos requisitos do utilizador, existindo assim vários tipos. Por exemplo, as métricas no *Google Analytics* incluem números comuns de *web analytics* como: visitas, páginas vistas, taxa de rejeição, taxa de conversão, receita [Cutroni, 2010], quantas vezes os ficheiros são acedidos, quantos endereços IP únicos acedem ao *site*, entre outros. Outra listagem de métricas pode-se encontrar no *Google Webmaster Tool* e no *Bing Webmaster Tool* tais como: impressões³², cliques³³, páginas rastreadas³⁴ e páginas indexadas³⁵. De todas estas métricas mencionadas, algumas têm mais impacto do que outras sobre um *site*. Listar quais têm mais importância, não faz sentido, por isso, tal dependerá do tipo de *site* que se está a medir. Se o *site* só vai exibir conteúdo, o conjunto de métricas não é relevante, mas se o *site* se dedica ao comércio de produtos, há um conjunto de métricas que se devem considerar [Ledford et al., 2010].

³² Número de vezes que o URL de um *site* aparece nos resultados de busca visualizados por um utilizador, excluindo as impressões de busca pagos do *AdWords* (o principal serviço de publicidade que o Google utiliza).

³³ Número de cliques nos URLs de um *site* nas SERP, excluindo os cliques nos resultados de pesquisa pagos do *AdWords*.

³⁴ Número de páginas verificadas pelos crawlers dos motores de busca.

³⁵ Número de páginas armazenadas no índice da base de dados dos motores de busca.

2.6. Ferramenta de Apoio para os *Webmasters*

Existe uma grande variedade de ferramentas para apoiar o trabalho dos *webmasters*, porém nesta seção faz-se referência apenas a duas: *Google Webmaster Tools* e *Bing Webmaster Tools*, por serem sugeridas pelos motores de busca mais conhecidos, serem gratuitas e por serem as mais usadas pelos *webmasters*. Ambas providenciam informações valiosas sobre como os motores de busca olham para um *site* incluindo: *links* de dados externos, *links* de dados internos, erros de rastreamento e termos de busca de alto volume [Enge et al. 2009].

2.6.1. *Google Webmaster Tools*

Google Webmaster Tools é um bom recurso para *webmasters* uma vez que com esta ferramenta é possível identificar facilmente um grande conjunto de problemas. Para se aceder a esta aplicação é preciso ter uma conta no Google. As suas características estão divididas em várias seções [MacDonald 2011], designadamente:

- **Visão geral.** Esta seção informa se o Google tem visitado o *site*, e se o *site* foi adicionado com sucesso ao seu índice.
- **Diagnóstico.** Esta seção avisa sobre quaisquer problemas que o Google encontrou, como meta dados incorretos ou páginas que não podem ter acesos.
- **Estatísticas.** Esta seção fornece informações sobre as buscas que levam os visitantes através dos motores de busca para o *site*. No entanto, podem-se obter estatísticas mais detalhadas usando o serviço de rastreamento do *Google Analytics*.
- **Links.** Esta seção é de grande ajuda para visualizar *links* externos, uma vez que exhibe uma listagem de *sites* que ligam ao *site* de origem.
- **Sitemaps.** Esta seção ajuda a construir *sitemaps*, um ficheiro especial que descreve a estrutura de um *site* e os seus ficheiros constituintes. O *sitemap* pode ser enviado para o Google e para os outros motores de busca de modo que eles saibam que páginas indexar.
- **Ferramentas.** Esta seção permite ajustar algumas configurações diversas do Google.

2.6.2. Bing Webmaster Tools

A Microsoft também oferece uma ferramenta para *webmasters* chamada *Bing Webmaster Tools* com serviços similares ao *Google Webmaster Tools*. Esta ferramenta é uma grande ajuda para se aferir o desempenho de um *site* e o tráfego que se está a obter. Recentemente, a Microsoft tem desenvolvido muitos esforços para melhorar a ferramenta continuando com atualizações novas e úteis [Chaters 2011]. A seguir apresentam-se algumas das suas características:

- **Painel.** Fornece uma visão geral e um resumo do que se encontra em cada uma das outras seções. A maior parte da informação que se precisa do Bing é acessível a partir deste painel.
- **Rastreamento.** O relatório de resumo Rastreamento apresenta a taxa sobre o *site* que está a ser rastreado e lista todos os erros de rastreamento que o *spider Bingbot* tem. Além disso, o separador de rastreamento permite ver detalhes do conteúdo que está a ser rastreado e o número de URLs submetidos, fornecendo acesso aos *sitemaps XML* que são detetados.
- **Índice.** O relatório de resumo Índice expõe quantas páginas foram indexadas pelo *Bingbot*.
- **Tráfego.** O relatório de resumo Tráfego apresenta informação valiosa sobre o tráfego na *web* com o qual se pode analisar o desempenho das consultas de busca ao longo do tempo.

2.7. Considerações Finais

Neste capítulo realizou-se a revisão bibliográfica sobre o funcionamento dos motores de busca, as técnicas SEO, bem como algumas ferramentas de apoio mais conhecidas que os *webmasters* utilizam em todo este âmbito do SEO.

Como pode ser visto, o objetivo dos motores de busca é dar aos seus utilizadores respostas relevantes, no menor tempo possível, e com conteúdos novos e relevantes, pelo que estes se encontram em constante mudança. Por outro lado, os proprietários de *sites* tentam otimizar os seus *sites* utilizando as técnicas SEO para estes serem

rastreados pelos motores de busca e conseqüentemente serem encontrados pelos utilizadores. A forma mais comum de verificar a performance destes *sites* é através das ferramentas *webmasters* e *analytics*, que além disso proporcionam dados estatísticos relevantes para melhorar o *site*. Todo este processo de otimizar páginas *web* faz-nos ver que as técnicas SEO estão em constante evolução, uma vez que estão sempre a surgir novas formas de otimização e os algoritmos dos motores de busca estão constantemente a ser aperfeiçoados.

Inicialmente os *sites* precisavam ser submetidos pelos *webmasters* aos motores de busca, por forma a serem rastreados pelos *crawlers* e colocados no índice das suas bases de dados [Pinkerton 2000]. Porém com o tempo, isto mudou e chegou a fazer-se automaticamente, graças ao aperfeiçoamento dos seus algoritmos. No início, a otimização para os motores de pesquisa dependia de informações fornecidas pelos *webmasters*, enquanto as etiquetas *meta keywords*, *meta description* e *title* eram as mais relevantes por posicionarem os *sites* no topo das SERP. Visto isto, os responsáveis de *sites* começaram a manipular o código HTML e o seu conteúdo para obter resultados mais otimizados nas SERP. Esta manipulação não foi bem sucedida e dividiu aos praticantes SEO em dois grupos, os praticantes que executam as técnicas *Black Hat* e os praticantes que executam as técnicas *White Hat*. Por esta razão, os motores de busca têm criado um conjunto claro de diretrizes para as práticas SEO, a fim de reduzir as técnicas *Black Hat* SEO [McDonald e Zayat, 2009].

Quanto às ferramentas de suporte para o trabalho SEO, entre pagas e gratuitas, existe um grande número no mercado. As mencionadas neste capítulo dão mais ênfase às ferramentas com que os *webmasters* desenvolvem os *sites* atualmente, e algumas delas são promovidas pelos próprios motores de busca. Cada ferramenta é usada segundo o critério do *webdesigner* e dos requisitos do *site*. Elas vão desde complementos para *browsers* como o SEOquake³⁶ a robustas ferramentas de análise como o Google Analytics, que são úteis para o posicionamento dos *sites*.

A existência de técnicas *Black Hat* SEO faz-nos perceber que as técnicas SEO éticas têm sido mal-usadas. Uma avaliação destas técnicas SEO éticas iria determinar as

³⁶ É um complemento de *Firefox* que fornece informação relevante de um *site*, como: páginas indexadas, *links* internos, *links* externos, *PageRank*, idade do domínio, entre outros.

mais eficazes. Estudos prévios como o artigo *Analysing Google Rankings through Search Engine Optimization Data* de Evans [2007], que identifica as técnicas mais populares de SEO através da análise de 50 *sites* e *Worst practices in search engine optimization* de Malaga [2008], que faz referência às técnicas *Black Hat* SEO, são basilares para a realização deste estudo. A tese *Model for Auditing Search Engine Optimization for E-business* de Schooner [2010] que avalia um *site* para verificar algumas técnicas SEO, também é fundamental para este estudo.

Outros estudos como *A comparative Study of Six Search Engines* de Verónis [2006] que fazem uma comparação entre os motores de busca Google, Yahoo e MSN e Voulà e Dir para determinar a relevância dos resultados, foram complementares para o estudo. Estes dois últimos motores de busca foram desenvolvidos por France Telecom e Iliad Group respetivamente. *Improving website visibility and information retrieval of e-commerce ventures: A specification to please the crawlers* de [Weideman et al. 2006] é outro estudo muito interessante que identifica elementos potenciais que providenciam o incremento da visibilidade nos *sites*. Para isto foram analisados 144 *websites* onde se verificou a estrutura dos *sites* e os fatores que impediam que se posicionassem nos motores de busca. *Correlation Analysis of On-Page Attributes and Search Engine Rankings* de Fister [2007] é uma tese que se centra exclusivamente em examinar a relação entre os atributos *on-page* e a classificação dos resultados de busca. Propõe um correlacionamento entre os atributos para obter elevadas classificações nos motores de busca e sugere aos fornecedores de conteúdo incluir *keywords* nas diferentes *tags* de um *sites*. Estes e outros, são estudos que fundamentaram a base da pesquisa que foi realizado.

Capítulo 3 - Descrição do Trabalho Efetuado

3.1. Visão Global do Trabalho

O trabalho efetuado sobre a avaliação das técnicas SEO está resumido na Ilustração 7. Atualmente as pessoas que possuem um negócio ou mesmo as empresas que comercializam produtos ou prestam serviços olham para a *web* como um negócio rentável onde podem aproximar-se do seu público-alvo. É por isso que uma grande maioria delas possui um *site* quer seja para se promover, para capturar potenciais clientes ou, simplesmente, para andar a par das novas tecnologias. Neste sentido, as pessoas e empresas de negócio tentam conseguir visibilidade dos seus *sites* nas SERP dos principais motores de busca tais como Google, Yahoo ou Bing.

O estudo da avaliação das técnicas SEO consta de três processos: o primeiro começa com a recolha de dados que permite ter informação relevante acerca do tema, o segundo é o processo de implementação onde se verifica o desempenho dos *sites* a serem analisados e a implementação das técnicas sobre eles. A verificação é feita através de ferramentas para *webmasters* e a implementação é monitorizada através de ferramentas *analytics*. No último processo, nomeado de teste, é criado um *site* com o objetivo de ver o comportamento das técnicas SEO desde a sua concepção. As vantagens que se podem obter deste estudo focam-se sobre três realidades: *sites*, motores de busca e utilizadores. O conhecimento das técnicas SEO mais eficazes permitiria aos *sites* obter a visibilidade que procuram nas SERP, visibilidade em um sentido de popularidade e reconhecimento, que atraia maior número de visitas. Por outro lado, a implementação destas técnicas irá ajudar os motores de busca na indexação de *sites*, sendo que neste sentido ambos serão beneficiados porque os *sites* que respeitam as diretrizes dos motores de busca, contribuirão para que os *crawlers* os rastreiem de uma maneira mais simples. Finalmente, o bom uso das técnicas SEO nos *sites* irá fornecer informação relevante aos utilizadores.

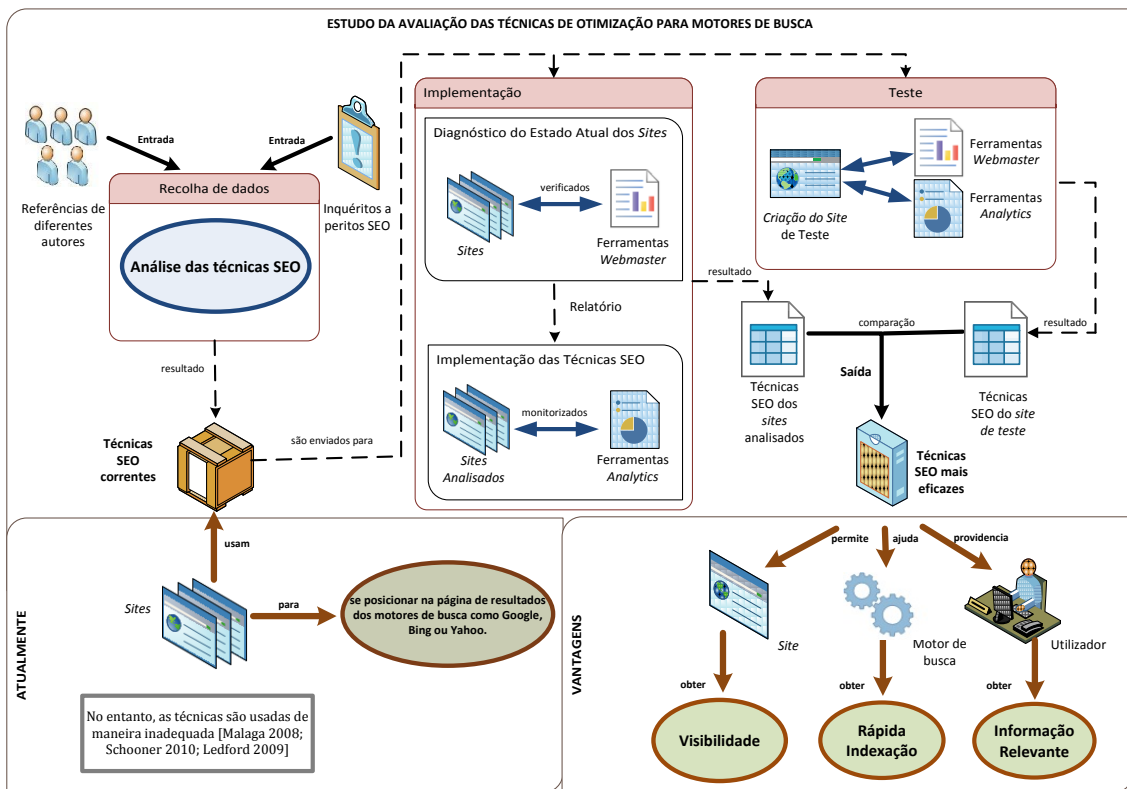


Ilustração 7. Processo do Estudo Efetuado. Fonte Própria.

3.1.1. Recolha de Dados

O estudo da avaliação das técnicas SEO começou com a pesquisa bibliográfica acerca da SEO, a fim de conhecer as técnicas correntes. Num início teve que se pesquisar diferentes fontes bibliográficas que apresentavam técnicas de otimização de *sites*, desde o ponto de vista técnico ao ponto de vista empresarial. Autores como Jones [2008], David [2011], [McDonald e Zayat 2009] e Thurow [2007] apresentam as técnicas SEO como as melhores práticas de otimização de um *site*, que atualmente fazem parte da estratégia de *Marketing Online*. Inicialmente SEO foi vista como um aspeto meramente técnico, porém ao longo do tempo com o surgimento dos *blogs*, *wikis*, e redes sociais, tornaram-se parte complementar da estratégia de *marketing* das empresas. Outra contribuição importante foi avançada por Malaga [2008], Marie [2012], Kent [2008] e Schimmel [2010] que apresentaram o outro lado das técnicas SEO, fazendo referência às práticas proibidas pelos motores de busca. Fontes como estas contribuíram para o conhecimento de técnicas SEO correntes. Igualmente, as entrevistas a diferentes peritos SEO contribuíram para o estudo pela partilha das suas

experiências. Estas duas entradas contribuíram à análise das técnicas e originaram uma tabela com as técnicas SEO mais correntes (Tabela 2), que se tem usado nestes últimos anos. Algumas delas ainda estão a ser usadas, enquanto outras já são consideradas como proibidas pelos motores de busca. Estas técnicas têm sido agrupadas em três categorias.

Categorias	Técnicas <i>White Hat</i> SEO	Técnicas <i>Black Hat</i> SEO
Conteúdo	<ul style="list-style-type: none"> • Conteúdo Próprio • Comunidades • RSS Feeds • <i>Linkbaiting</i> 	<ul style="list-style-type: none"> • Conteúdo Oculto • <i>Scraping Content</i> • <i>Article Spinning</i>
Código	<ul style="list-style-type: none"> • Linguagem de Estilo • URLs curtas • <i>Tagging</i> • Criação de <i>sitemap</i> XML 	<ul style="list-style-type: none"> • <i>Doorway Pages</i> • <i>Cloaking</i> • <i>Keywords Stuffing</i> • Páginas PPC
Ligações	<ul style="list-style-type: none"> • <i>Linkbuilding</i> • Submissão a Diretorias • Redes sociais • <i>Article submission</i> • <i>Bookmarking Social</i> 	<ul style="list-style-type: none"> • <i>Links</i> enganosos • <i>Link Farms</i> • <i>Wiki Spam</i> • <i>Blog/Forum Spam</i>

Tabela 2. Conjunto de Técnicas SEO. Fonte: Própria.

3.1.2. Implementação

Nesta fase realiza-se a verificação de três *sites* com diferentes temáticas e a implementação das técnicas *White Hat* SEO sobre estes, por forma a fazer um seguimento da evolução dos *sites* com as técnicas SEO implementadas. Este processo está dividido em duas fases: Verificação e Implementação, ambos são apresentados em seguida:

Verificação

Depois de se ter a tabela com as técnicas SEO mais recentes, iniciou-se a fase de verificação dos *sites*. Nesta fase usaram-se três *sites* de diferentes temáticas. Devido à privacidade de informação, os nomes dos domínios serão reservados e subsequentemente serão nomeados como *site A*, *site B* e *site C*. O *site A* pertence ao

setor turismo, o *site B* pertence ao setor de tecnologias e o *site C* pertence ao setor legislativo. A experiência da autora na área de posicionamento de *sites* facilitou o contacto via *email* com os proprietários dos *sites*, uma vez que foram clientes de projetos anteriores. A escolha dos *sites* fez-se de maneira aleatória sendo que o único critério considerado foi a aprovação pelos proprietários, de disponibilizar os seus *sites* como parte da avaliação durante o estudo. A principal preocupação dos proprietários, ao disponibilizar precisamente aqueles *sites*, era conseguir tráfego que gerasse lucro para os seus negócios.

A fase de verificação consiste em diagnosticar o estado inicial dos *sites*, ou seja, como é que eles estão a funcionar, se eles estão a usar técnicas SEO, se eles estão a cumprir com as diretrizes dos motores de busca e, quais são os possíveis erros que o motor de busca detetou através de uma ferramenta de apoio para os *Webmasters*. Neste caso foram usadas o Google Webmaster Tool, Bing Webmaster Tool, o Link Checker da W3C e o complemento SEOquake. Estas duas últimas, embora não tenham sido mencionadas na revisão de literatura, foram usadas como suporte à informação contida nas ferramentas para *webmasters*, de modo a verificar os erros nos *links*.

Para agilizar o trabalho de verificação dos *sites* foram criados modelos de relatórios que se consideraram relevantes. O primeiro está focado na visão geral do *site* e o segundo no estado inicial do *site*, sendo os mesmos apresentados seguidamente.

Visão Geral do Site

O modelo do relatório da “Visão Geral do *Site*” apresenta uma *descrição* do *site* que está a ser verificado, através do *nome* e do *setor* ao qual pertence. Ainda fornece informação sobre a idade, *pagerank* e a data do último acesso ao *site*. A verificação do *site* fez-se com base na página principal, mas o relatório possui um campo com duas alternativas *página principal* e *outras páginas*, a última alternativa é utilizada caso se verifique as outras páginas do *site*. O modelo divide-se em quatro colunas: elemento, estado, descrição e observações, de modo a ter uma orientação geral do *site*. No campo *elemento*, encontram-se os componentes mais básicos que todo diagnóstico de *site* deve considerar; no campo *estado*, situa-se uma valoração qualitativa da forma como é percebido o componente no *site*, variando entre *bom*, *regular* e *mau*; no campo

descrição, descreve-se cada componente encontrado no *site*; e no campo *observações*, salientam-se aspetos que o *site* tem de melhorar. A seguir apresenta-se a visão geral de cada *site* diagnosticado (Relatórios 1, 2 e 3), segundo o modelo descrito.

Visão Geral do <i>Site</i>			
Descrição do Site	Nome: <i>Site A</i>	Setor: Turismo	
	Idade do domínio: 3 anos e 5 meses	PageRank: 2	
Último acesso	31-08-2012		
Página de análise	Página Principal <input checked="" type="checkbox"/>	Outra página _____ <input type="checkbox"/>	
Elemento	Descrição	Estado	Observações
URL	Palavra de 10 caracteres	Regular	Sugere-se que o nome de domínio seja uma palavra-chave do setor ao invés do nome da empresa.
<i>Title Tag</i>	24 caracteres	Regular	O título pode ser melhorado adicionando-se palavras-chave.
<i>Meta Description</i>	170 caracteres	Regular	Sugere-se colocar palavras entre 70 e 160 caracteres
<i>Meta Keywords</i>	Palavras-chave: 28 Carateres: 426	Mau	Palavras-chave em excesso, no entanto, não irá afetar a classificação nas SERP, porquanto o <i>tag</i> já não é considerado relevante.
<i>Headings</i>	H1, H2, H3, H4, H5 e H6 encontrados: 0	Mau	Sugere-se criar cabeçalhos HTML onde se coloque palavras-chave que otimizem a página.
Conteúdo	Texto útil	Regular	Sugere-se adicionar palavras-chave nos textos e ordena-los de modo que andem a par com os títulos das páginas.

Imagens	Imagens encontradas: 9 Atributos ALT todos preenchidos	Bom	
Frames	Frames encontrados: 0	Bom	
Flash	Não	Bom	

Relatório 1. Visão Global do Site A.

Visão Geral do Site			
Descrição do site	Nome: Site B		Setor: Tecnologias
	Idade do domínio: 3 anos e 7 meses		PageRank: 1
Último acesso	31-08-2012		
Página de análise	Página Principal <input checked="" type="checkbox"/>	Outra página _____ <input type="checkbox"/>	
Elemento	Descrição	Estado	Observações
URL	Palavra de 11 caracteres	Bom	
Title Tag	Expressão de 12 palavras com 78 caracteres	Regular	Sugere-se dar espaços entre as palavras e rever ortografia, uma vez que os <i>crawlers</i> reconhecem-no como se fossem 10 palavras.
Meta Description	142 caracteres	Bom	
Meta Keywords	Palavras-chave: 10 Carateres: 134	Bom	

<i>Headings</i>	H2 encontrados: 5 H1, H3, H4, H5 e H6 encontrados: 0	Regular	Sugere-se criar cabeçalhos HTML onde se coloquem palavras-chave que otimizem a página.
Conteúdo	Texto básico	Regular	Sugere-se adicionar mais texto no site de modo que atraia os utilizadores.
Imagens	Imagens encontradas: 62 Atributos ALT sem uso: 48	Mau	Requer revisão e sugere-se adicionar o atributo ALT em todas as imagens.
<i>Frames</i>	<i>Frames</i> encontrados: 0	Bom	
<i>Flash</i>	Detetado	Regular	Sugere-se não usar Flash em todo o <i>site</i> , uma vez que pode ser um dos factores que impede a classificação nos motores de busca. Pode-se usar apenas nos botões de navegação, e na <i>homepage</i> colocar apenas imagens ligeiras que não provoquem lentidão no <i>site</i> .

Relatório 2. Visão global do *Site B*.

Visão Geral do <i>Site</i>			
Descrição do <i>site</i>	Nome: <i>Site C</i>		Setor: Legislativo
	Idade do domínio: 6 meses		PageRank: 0
Último acesso	31-08-2012		
Página de análise	Página Principal <input checked="" type="checkbox"/>		Outra página _____ <input type="checkbox"/>
Elemento	Descrição	Estado	Observações
URL	Palavra de 13 caracteres	Bom	

Title Tag	101 caracteres	Mau	Sugere-se colocar o título entre 10 e 70 caracteres, incluindo espaços
Meta Description	354 caracteres	Mau	Sugere-se colocar palavras entre 70 e 155 caracteres
Meta Keywords	Palavras-chave: 12 Carateres:17	Bom	
Headings	H4 encontrados: 1 H1, H2, H3, H5 e H6 encontrados: 0	Regular	Sugere-se criar cabeçalhos HTML onde se coloquem palavras-chave que otimizem a página.
Conteúdo	Texto básico	Regular	Sugere-se adicionar mais texto no <i>site</i> de modo que atraia os utilizadores.
Imagens	Imagens encontradas: 96 Atributos ALT sem uso: 88	Mau	Requer revisão e sugere-se adicionar o atributo ALT em todas as imagens.
Frames	Frames encontrados: 0	Bom	
Flash	Não	Regular	

Relatório 3. Visão Global do *Site C*.

Prestar atenção aos elementos apresentados nos relatórios é uma grande ajuda para verificar os fatores que impedem o posicionamento nas SERP, e resolver estes problemas ajuda a ter o *site* otimizado. Uma vez corrigidos os erros, o *site* ficará pronto para ser rastreado pelos *crawlers*. Não há dúvida que providenciar informação útil aos motores de busca irá impulsionar o tráfego do *site*. Por exemplo, o texto colocado nas etiquetas *title* e *meta description* muitas vezes aparece na listagem dos resultados de busca como o cartão-de-visita do *site*, portanto, fornecer um texto descritivo e útil aos motores de busca tem mais possibilidades de ser clicado pelos utilizadores.-

Estado Atual do Site

O modelo do relatório do “Estado Atual do Site” é muito parecido ao modelo de “Visão Geral do Site”, uma vez que possui uma descrição do *site*, com o seu respetivo nome, setor, idade, *pagerank* e data em que foi acedida. Os campos estão divididos em três seções, a primeira é a seção onde estão localizadas as principais categorias de análise, na segunda encontra-se cada elemento analisado, e a terceira são as observações que cada elemento apresenta. Os relatórios 4, 5 e 6 apresentam os estados atuais de cada *site* analisado.

Estado Atual do Site			
Informação do Site	Nome: <i>Site A</i>		Sector: Turismo
	Idade do domínio: 3 anos e 5 meses		PageRank: 2
Último acesso	31-08-2012		
Seções	Elementos de análise		Observações
Código HTML	<i>Erros on-page</i>	<i>Etiquetas Title</i> duplicadas	0
		<i>Etiquetas Meta Description</i> duplicadas	3
	<i>Sitemap XML</i>		<i>Não</i>
Rastreamento	Erros de rastreamento	Páginas não encontradas	4
		Aplicações não encontradas	1
		Robot.txt	<i>Não</i>
	URL Bloqueados		0
	Estado do Índice	Número de páginas indexadas	114
	Texto âncora		9
Tráfego	Consultas de Pesquisa		98

	Cliques	Impressões	
	70	4500	
	Links Externos	Outbound links	1
		Inbound links	7047
	Links Internos		26

Relatório 4. Estado Atual do Site A.

Estado Atual do Site			
Informação do Site	Nome: Site B		Sector: Tecnologias
	Idade do domínio: 3 anos e 7 meses		PageRank: 1
Último acesso	31-08-2012		
Seções	Elementos de Análise		Observações
Estrutura de site	Código HTML	Etiquetas Title duplicadas	7
		Etiquetas Meta Description duplicadas	18
	Sitemap		Não
Rastreamento	Erros de rastreamento	Páginas não encontradas	18
		Aplicações não encontradas	1
		Robot.txt	Não
	URL Bloqueados		0
	Texto âncora		10
Tráfego	Consultas de Pesquisa		13
	Cliques	Impressões	
	13	600	

	<i>Links Externos</i>	<i>Outbound links</i>	10
		<i>Inbound links</i>	4792
	<i>Links Internos</i>		12

Relatório 5. Estado Atual do *Site B*.

Estado Atual do <i>Site</i>			
Informação do <i>Site</i>	Nome: <i>Site C</i>		Sector: Legislativo
	Idade do domínio: 6 meses		PageRank: 0
Último acesso	31-08-2012		
Seções	Elemento de Análise		Observações
Estrutura de <i>site</i>	Código HTML	<i>Etiquetas Title</i> duplicadas	0
		<i>Etiquetas Meta Description</i> duplicadas	0
	<i>Sitemap</i>		1
Rastreamento	Erros de rastreamento	Páginas não encontradas	0
		Aplicações não encontradas	0
		Robot.txt	<i>Sim</i>
	URL Bloqueados		1
	Estado do Índice	Número de páginas indexadas	1
	Texto âncora		1
Tráfego	Consultas de Pesquisa		5
	Cliques	Impressões	
	35	400	
	<i>Links Externos</i>	<i>Outbound links</i>	4

	<i>Inbound links</i>	0
<i>Links Internos</i>		14

Relatório 6. Estado Atual do *Site C*.

Implementação

Depois de obter o relatório de verificação para cada *site*, iniciou-se a fase de implementação das técnicas SEO. Para prosseguir este trabalho de maneira ordenada e correta fez-se uso do processo SEO [Malaga 2008; McVittie 2012] apresentado no estado de arte. Uma vez implementadas as técnicas, estas foram monitorizadas pelo Google Analytics. Antes de prosseguir com as etapas do processo SEO, deve-se ter em conta que a implementação das técnicas realizou-se segundo as necessidades encontradas nos relatórios de Verificação. Pelo que não se utilizaram todas as técnicas *White Hat* SEO apresentadas na tabela 2 em cada *site*. Cada uma delas é independente.

- 1. Primeira etapa:** Para o *site A, B e C* pesquisaram-se palavras-chave relevantes que acrescentem valor ao negócio. A ferramenta *Google Keywords Sugesstion Tool* foi bastante útil nesta etapa.
- 2. Segunda etapa:** Na segunda etapa corrigiram-se erros, acrescentaram-se alguns elementos, e alteraram-se outros, segundo as observações obtidas dos relatórios de análise. Na ilustração 8 providencia-se um exemplo dessa alteração.

```

11 <head>
12 <meta name="msvalidate.01" content="3207C73DB58FFE64C33AD38DF2EB033" />
13 <meta name="google-site-verification" content="Kb_a3VzVemw8cpv9_25KHFvc_Qjsn0slqh9f130ZUw8" />
14 <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
15 <title>
16 <META NAME="Keywords" CONTENT="peru travel, travel cuzco, trip lima, tour puno, cuzco tour, inka travel, machupicchu travel, travel to machu picchu, cuzco mac
17 picchu, tour machu picchu, machu picchu tours, cuzco tours, travel south america, photos cuzco, peru restaurants, cuzco peru, sacred valley, photos cuzco,
18 museums cuzco, restaurants cuzco, peru photos, lima tours, maldonado, puno tours, ica nazca lines, train tickets, peru information, peru news">
19 <META NAME="Description" CONTENT="Destinations Peru, tours to Machu Picchu, Inca Trail, Nazca Lines, Titicaca Lake, Manu National Park , Tambopata Reserve, Co
20 Canyon, Pacaya Samiria, hotel reservations">
21 <META NAME="Classification" CONTENT="Travel">
22 <META NAME="Abstract" CONTENT="Peru Adventure & Cultural Tour Operator">

```

Ilustração 8. Alterando os Tags de *Site A*.

- 3. Terceira etapa:** Uma vez feitas as otimizações, iniciou-se com a promoção do *site*, isto é, obter tráfego para o *site*. Inicialmente, verificou-se a posição de cada

site nas SERP com a palavra-chave que se pretendia promover, por exemplo o *site* A foi confirmada para estas: Travel Peru, Tour Peru, Travel Machupicchu, das quais nenhuma apareceu dentro dos primeiros cinquenta *sites* nos resultados de busca dos motores de pesquisa já conhecidos. A verificação foi realizada a fim de se ter uma noção da concorrência e do tempo aproximado que poderia levar a cabo todo este processo. Subsequentemente, fez-se uso da técnica de submissão a diretorias, que é uma técnica SEO muito usada pela maioria de *Webmasters* ao iniciarem a promoção de um *site*. A finalidade desta técnica é obter *backlinks*, que são um meio fundamental para a classificação nos motores de busca. Quanto mais *backlinks* tenha o *site*, maior é a hipótese de se posicionar nas primeiras páginas dos resultados de busca.

- 4. Quarta etapa:** A última etapa foi fazer um seguimento do trabalho anterior, ou seja, monitorizar a evolução dos *sites* enquanto era utilizada a técnica SEO. A monitorização constante de tráfego e a apresentação de relatórios foram realizadas através de Google Analytics. Os resultados do tráfego obtidos durante este procedimento apresentam-se no capítulo seguinte.

3.1.3. Teste

Está é o processo final do estudo da avaliação das técnicas SEO. Neste processo é criado um *site* de teste, onde se faz um seguimento as técnicas *White Hat* SEO desde a conceção de um *site*. O trabalho foi dividido em dois grupos: *Mock-up* do *Site* e *Criação* e Monitorização do *Site*. A seguir apresentam-se estas duas categorias:

Mock-up do Site

A estrutura de um *site* não é tão simples como parece, muitas vezes é complexa e faz com que uma página *web* não obtenha a qualidade necessária, não sendo por isso encontrada pelos utilizadores nos resultados de busca. Uma boa forma de estruturar um *site* é fazendo um *mock-up*, que permita mostrar as funcionalidades e características que o *site* irá suportar. Neste sentido a estruturação do *site* de teste considerou diversos

critérios como: as orientações dos motores de busca, a experiência da autora e a sugestão de alguns autores como Miller [2000], que determina a qualidade de um *site* através de diferentes fatores. Em seguida são apresentados esses fatores:

- **Browser.** Devido à existência de diversos tipos de *browsers*, um *site* deve ser desenvolvido só com elementos que funcionem na maioria deles.
- **Tecnologias de Visualização.** O que se observa no *browser* é um *site* desenvolvido por diferentes tecnologias, aplicações ou pacotes de software, como HTML, Java, JavaScript, ActiveX, Cgi-Bin Scripts, Database Access, entre outros, que dependendo do tipo de *site* que se está a desenvolver devem ser facilmente comprováveis e tratados de modo que o *site* os suporte.
- **Navegação.** Os utilizadores navegam pelo *site*, indo de uma página para outra, fazendo cliques nos *links*, imagens, entre outros, pelo que a navegação tem que ser rápida e sem erros.
- **Modo objeto.** Os campos, botões de entrada, botões de opção, caixas de seleção, formulários, etc., devem ser tratados em modo objeto, ou seja, independentemente das fontes e preferências, de modo que as validações de qualidade sejam independentes da representação.
- **Resposta do servidor.** É aqui onde se verifica a velocidade com que o *web hosting* responde ao utilizador enquanto ele navega no *site*. A carga da *internet* também afeta a resposta, no entanto esse fator está geralmente fora do controle do *webmaster*. A capacidade do *hardware* do servidor e a taxa de transferência são os problemas mais comuns, pelo que devem ser levados em consideração.
- **Interação e *feedback*.** Para *sites* com conteúdo estático, o único problema é a disponibilidade mas para *sites* que interagem com o utilizador, a velocidade e a confiabilidade com as quais se interagem podem ser grandes fatores, pelo que se sugere que o *feedback* seja rápido e que gere confiança no utilizador.
- **Utilizadores simultâneos.** Enquanto os *sites* frequentemente se assemelham a estruturas de *software* convencionais cliente/servidor com vários utilizadores em vários pontos, um *site* pode ser muito diferente e muito mais complexo que aplicações complexas, pelo que se deve prestar atenção ao número de utilizadores que interagem e a forma como eles se contactam entre si através do *site*.

Estes elementos foram levados em conta para estruturar o *site* de teste. A seguir apresenta-se o *mock-up* final do *site* (Ilustração 9).



Ilustração 9. *Mock-up* do *Site* de Teste.

Criação do *Site* de Teste

O *site* de teste é um *site* onde as técnicas *White Hat* SEO são avaliadas, desde a conceção do *site*. Antes da criação do *site* analisaram-se vários critérios, podendo dizer-se que foi realizado um estudo de mercado³⁷. O estudo incluiu três análises importantes: Análise de clientes, análise da concorrência e estratégia [Burgemeister 2003].

- 1. Análise de clientes.** Neste sentido, analisou-se a tendência das buscas mais frequentes na *web* pelos utilizadores dos motores de busca, através do *Google Trends*³⁸. Geralmente, as áreas com maior interesse na criação de *sites* são as de

³⁷ Frequentemente o termo estudo de mercado é usado para descrever o processo de identificar e recolher informação relacionada a um produto ou serviço específico [Nykiel 2007].

³⁸ É uma ferramenta que mostra com que frequência se realiza uma busca específica em várias regiões do mundo e em várias línguas.

negócio, educação e viagens, pelo facto de serem as mais procuradas na *web*. Entre as três áreas, a Ilustração 10 dá-nos uma noção de que a busca mais frequente foca-se na área de Educação. Foi assim que se concebeu a ideia de fazer um *site* para a área de Educação. Os termos de pesquisa para ver a tendência das áreas foram introduzidos na língua inglesa por ser a língua com maior número de utilizadores na *web* [IWS 2012].

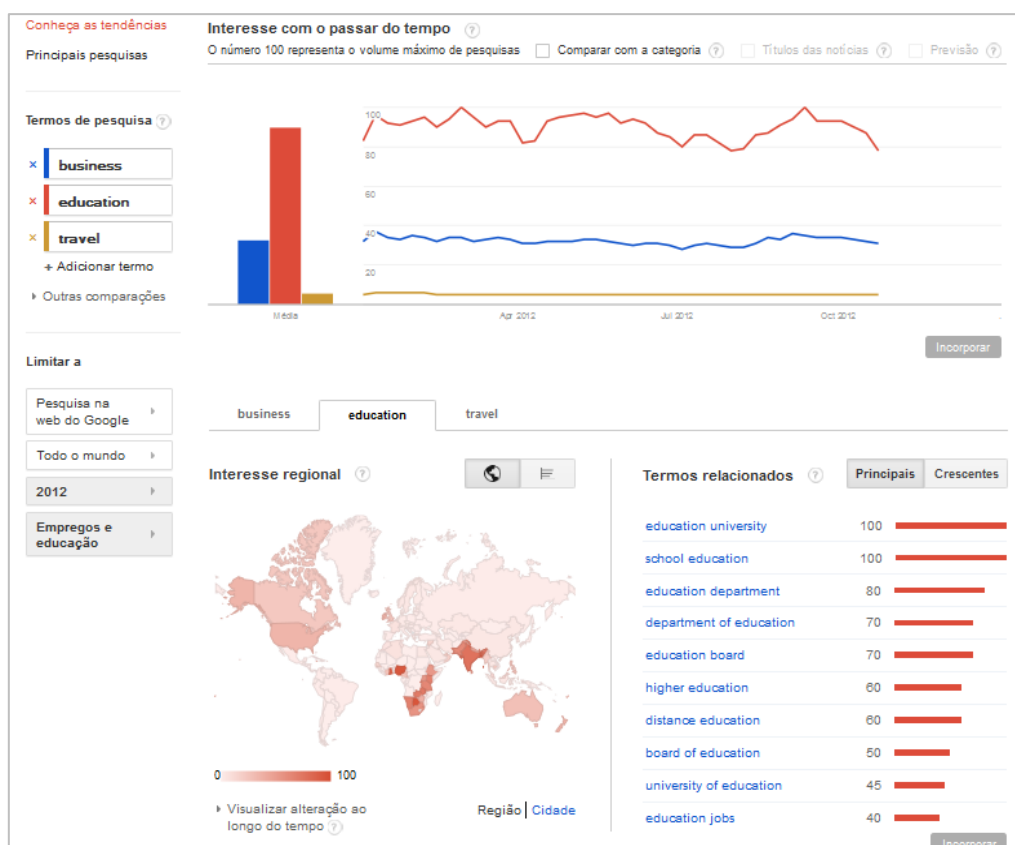


Ilustração 10. Tendência de Busca a Nivel Mundial para as Áreas de Negócio, Educação e Viagens no Ano 2012.

Consequentemente teve de se determinar uma subcategoria mais específica no que respeita à Educação. O facto de ter participado em um programa Erasmus incentivou a autora a focar o *site* para esta subcategoria, uma vez que viu a necessidade de obter informação relevante antes e durante o período de estudo. Da mesma forma que para as áreas anteriores (negócio, educação e viagens), se recorreu ao *Google Trends* para analisar a tendência do termo de busca “Erasmus”, verificou-se que o termo “Erasmus Mundus” é o mais pesquisado

(Ilustração 11). Motivo pelo qual decidiu-se criar um *site* destinado aos estudantes *Erasmus*.

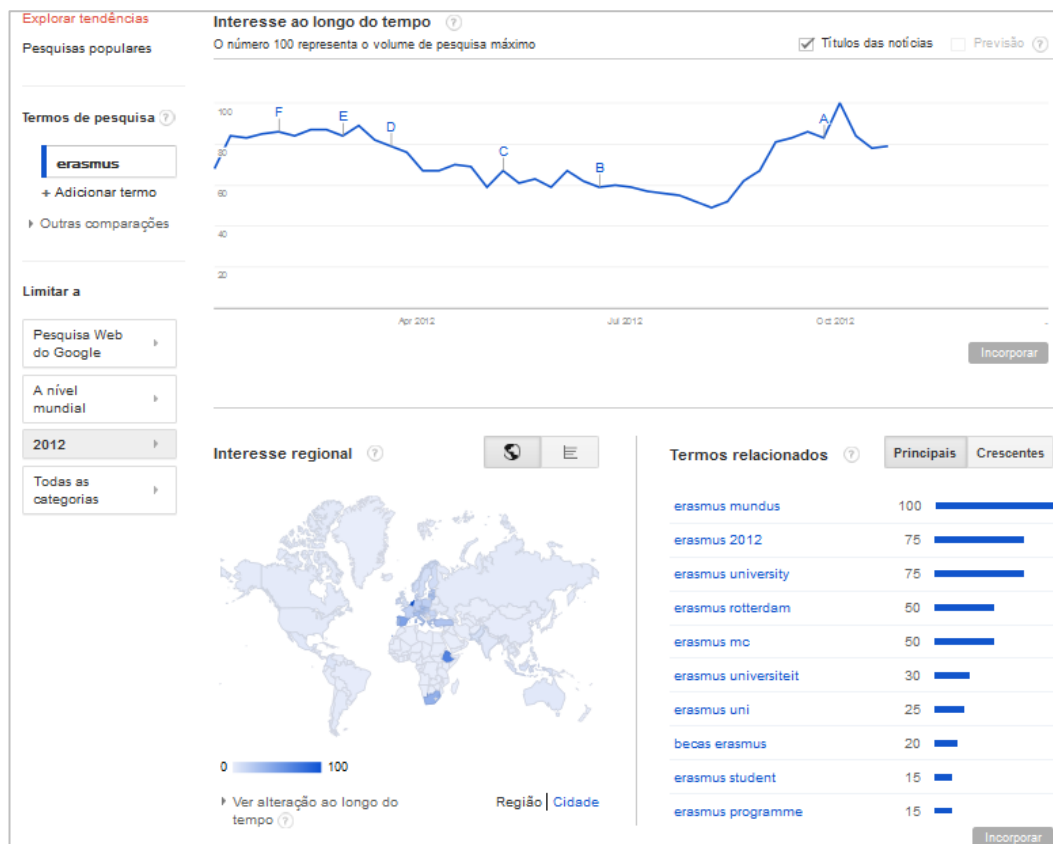
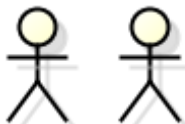


Ilustração 11. Tendência de Busca a Nível Mundial para o Termo *Erasmus* no Ano 2012

2. Análise da concorrência. A concorrência para os *sites Erasmus* não é grande, mas existem *sites* específicos que providenciam informação do programa, como o *site* oficial do *Erasmus Mundus*. Além disso, encontrou-se um potencial concorrente que fornece informação relevante aos alunos *Erasmus*, no entanto o *site* funciona como um fórum. Um *site* que funciona como fórum geralmente é gerida por varios gestores de conteúdo, uma vez que as perguntas e respostas devem ser verificadas constantemente para não perder de vista o utilizador. Neste sentido, o que se pretende fazer com o *site* de teste é providenciar informação relevante das experiencias Erasmus com alunos que tenham participado no programa, os quais possam partilhar a suas experiencias aos novos ou futuros alunos de Erasmus através do *site*, de modo que não se inibam de iniciar um programa como este, nem tenham dificuldade durante ou antes do periodo de estudo.

- 3. Estratégia.** Este item está baseado na estrutura do *site* e nos elementos que irão diferenciá-lo dos seus concorrentes. Em seguida é apresentado o esquema da estratégia:

Publico alvo



Estudantes *Erasmus*

(Estudantes *Erasmus* Europeus e Estudantes *Erasmus Mundus*)

Divisão do *site*

- a) Utilizador
- b) Universidades
- c) Cidades

a) Utilizador

A página nomeada *Erasmus* foi criada para dar acesso aos estudantes *Erasmus*. O utilizador *Erasmus* poderá registar-se e iniciar sessão para preencher a sua experiência e interagir com os outros utilizadores, através de perguntas ou comentários sobre os perfis que lhes interesse. Também podem procurar informação sobre as cidades e universidades às quais estão destinados para fazer Erasmus.

b) Universidades

Neste item se encontrará informação relevante sobre as universidades.

c) Cidades

Neste outro item poderá encontrar-se informação referente as cidades *Erasmus*, por exemplo uma descrição da cidade, a sua população, alguma história, entre outros.

Uma vez feita a análise do mercado prosseguiu-se com a implementação do *site* em conjunto com as técnicas *White Hat* SEO listadas na Tabela 2, além de se ter em consideração as diretrizes dos motores de busca. O processo SEO [Malaga 2008; McVittie 2012] tido em conta apresenta-se a seguir:

Primeira Etapa: Nesta primeira etapa escolheu-se o nome do domínio com base nas análises da concorrência existente na *web*. Esta etapa é similar à primeira etapa da implementação de técnicas SEO nos *sites* analisados (*site* A, B e C), só que neste caso a pesquisa de palavras-chave aplica-se também ao nome de domínio. Ascensão [2011] providencia uma explicação mais detalhada sobre a pesquisa de palavras-chave antes de começar a criar o *site* e recolher informação para o conteúdo. O *Google Keywords Tool* serviu de apoio para esta etapa (Ilustração 12).

The screenshot shows the Google Keyword Tool interface. At the top, there are options for 'Download', 'View as text', and 'Sorted by Relevance'. Below this, there are two sections: 'Search terms (1)' and 'Keyword ideas (9)'. Each section has a 'Save all' button and a search term input field. The 'Search terms (1)' section shows one result: 'erasmus experience' with a competition level of 'Low', 1,300 global monthly searches, and 16 local monthly searches. The 'Keyword ideas (9)' section shows nine suggestions with their respective competition levels, global monthly searches, and local monthly searches.

Keyword	Competition	Global Monthly Searches	Local Monthly Searches
erasmus experience	Low	1,300	16
Keyword ideas (9)			
erasmus student	Low	9,900	320
erasmus year	Low	1,000	210
erasmus scheme	Low	6,600	140
erasmus study	Low	2,400	210
erasmus experiences	Low	880	< 10
erasmus programe	Low	880	12
the erasmus program	Low	27,100	880
the erasmus programme	Low	18,100	390
erasmus statistics	Low	320	16

Ilustração 12. Escolha do Nome de Domínio do *Site* de Prova.

Uma vez que o *site* de teste é destinado a apresentar experiências dos estudantes *Erasmus*, optou-se pelo nome de domínio: *Erasmus Experiences*. Na ilustração 13 é visível que o termo *Erasmus Experiences* é uma das alternativas de palavras-chave do termo *Erasmus Experience* que o *Google Keywords Tool* sugere. No início optou-se por esse nome de domínio mas não se obteve a sua disponibilidade nos servidores de domínios, pelo que se escolheu o termo mais próximo: *Erasmus Experiences*. Na ilustração, também se nota que a concorrência é baixa, as buscas globais mensais³⁹ correspondem a 880 pesquisas por mês e as pesquisas locais mensais são menores a 10. Estes dados estatísticos podem não ser

³⁹ O número de vezes que é digitado uma palavra-chave na caixa do motor de busca.

muito animadores quanto ao tráfego, porém o nome é o mais adequado por estar alinhado com o que se pretende fazer com o *site*.

Segunda Etapa: A segunda etapa consistiu basicamente na criação do *site*, nomeadamente a codificação e a introdução do conteúdo. Nesta fase tentou-se tornar o *site* o mais amigável possível para os motores de busca, ou seja, o mais Search Engine Friendly (SEF) possível, de modo a ser rastreado de forma mais simples pelos *crawlers*.

Terceira Etapa: Esta etapa utilizou a técnica *bookmarking* social a fim de obter tráfego no *site*. Para além desta técnica se *planeou* promover o *site* através da rede social Facebook entre o grupo de alunos Erasmus como parte deste estudo, quem serão solicitados para preencher a suas próprias experiências.

Quarta Etapa: Esta etapa andou a par com a etapa anterior. A monitorização constante e a apresentação de relatórios foi uma forma de verificar como estava a evoluir o *site* enquanto eram adicionadas algumas funcionalidades ou implementada alguma técnica *White Hat* SEO. O *Google Analytics* fez parte desta monitorização. Os resultados são apresentados no capítulo a seguir.

3.3. Entrevistas

As entrevistas foram realizadas com base em um guião (Anexo II) previamente desenvolvido, envolvendo questões-chave para efeitos práticos deste estudo. Importa mencionar que por ser uma entrevista de tipo semiestruturado fizeram-se questões adicionais ao guião que foram adaptadas a cada um dos entrevistados.

Os peritos SEO envolvidos na entrevista foram cinco. Cada um deles deu uma percepção de SEO segundo a sua experiência e o trabalho que desenvolve na área. As entrevistas foram registadas em suporte áudio, e subsequentemente revistas e analisadas para extrair informação pontual do tema em estudo. As percepções dos peritos serveram como base para levar em consideração quais técnicas são consideradas eficazes, que ferramentas de análise são usadas no âmbito SEO, como define-se o termo SEO, e alguns comentários finais sobre a área em questão. O Anexo I apresenta a mensagem

eletrónica de pedido de entrevista e o Anexo III mostra um resumo das entrevistas realizadas.

As propostas iniciais para efetuar as entrevistas foram duas, uma presencial e outra *online*. Tentou-se fazer ambas, no entanto, pela disponibilidade dos entrevistados, optou-se pelo meio *online*. As entrevistas duraram uma média de 18 minutos por participante. O número de perguntas respondidas por participante foram 10 aproximadamente, dependendo do entrevistado, sendo no entanto mantida uma estrutura base das perguntas (Anexo II). As perguntas giravam à volta de SEO, motores de busca e ferramentas de apoio para otimizar *sites*. Na Tabela 5 se disponibiliza uma breve descrição dos peritos entrevistados:

Entrevistado	Setor da Empresa	Função/Experiencia	Data da entrevista
Entrevistado A	Educação	<i>Webmaster</i> /Técnico informático com quatro anos de experiência na área de SEO.	24-05-2012
Entrevistado B	Consultora de Informática, especializada em Marketing Online	Chefe de projetos e responsável do contacto com os clientes/ Engenheiro de sistemas com quatro anos de experiência na área SEO.	02-06-2012
Entrevistado C	Consultoria de negócios em tecnologias da informação. Agência de viagens.	<i>Webdesigner</i> /Informatico com oito anos de experiencia no <i>design</i> de <i>sites</i> . Atualmente é responsável por vários projetos em duas empresas.	04-06-2012
Entrevistado D	Informática, dedicada a criação e promoção de <i>sites</i> .	Gestor de equipa/Engenheiro de sistemas com cinco anos de experiência em SEO.	04-07-2012
Entrevistado E	Empresa dedicada ao <i>marketing online</i>	Gestor de projeto/Engenheiro de sistemas com seis anos de experiência na área de SEO.	15-07-2012

Tabela 3. Descrição dos Peritos Entrevistados

Capítulo 4 - Resultados

4.1. Resultados da Entrevista

As entrevistas foram realizadas em meados de 2012, durante o período de um mês e três semanas. O tipo de entrevista usado foi o semiestruturado e a interação com os intervenientes foi *online*, com uma duração aproximadamente de 18 minutos por participante, sendo cada uma delas foi registada em formato áudio. Os cinco entrevistados são peritos no âmbito SEO e o número de anos de experiência que possuem na área variam entre os quatro e os seis anos.

No decorrer das entrevistas, os participantes tentaram focar as suas respostas tendo por base o guião definido. Um resumo das entrevistas é apresentado no Anexo III. A seguir apresentam-se os resultados das entrevistas, focados na definição SEO e as técnicas SEO usadas atualmente:

O *Entrevistado A* define SEO como uma estratégia que permite aos *sites* aparecer na primeira página dos resultados de busca através da construção de *links*, redes sociais ou outras tecnologias. As técnicas SEO que atualmente usa nos seus *sites* são os *bookmarkings* sociais⁴⁰, rara vez *linkbuilding*, e na tentativa de manter os utilizadores com as últimas novidades, faz uso dos *blogs*. Menciona ainda que o conteúdo dinâmico e algumas alterações *on-page* são as técnicas mais eficazes.

O *Entrevistado B* define a SEO como um processo de otimização de elementos *on-page* e *off-page* de uma página para os motores de busca, particularmente para o Google. No que respeita as técnicas SEO, considera que uma boa estrutura de *links* internos, a criação de um *sitemap*, um bom conteúdo, a criação de *blogs*, e o intercâmbio de *links*, são as mais adequadas para a otimização dos seus *sites*.

O *Entrevistado C* define a SEO como uma forma de otimizar *sites* para os motores de busca e algo com que muitas pessoas estão a trabalhar recentemente. Quanto às técnicas SEO que usa nos seus *sites*, refere que um bom *design*, uma boa estrutura de

⁴⁰ São serviços *online* que permite aos internautas, armazenar, classificar, editar e partilhar os seus favoritos ou marcadores de documentos *web* através da *internet*. Os mais destacados são o *Delicious* e o *StumbleUpon*.

código fonte baseando-se em *standards*, a correta utilização de *tags* e palavras-chave são fundamentais para o posicionamento nos motores de busca.

O *Entrevistado D* define a SEO como o acrónimo de *Search Engine Optimization* e sublinha que se pode referir tanto à otimização de *sites*, para conseguir visibilidade no topo dos resultados dos motores de busca, bem como à pessoa responsável de fazer esse trabalho. As técnicas SEO que considera relevantes para os seus *sites* são: conteúdo relevante, submissão à diretórias, intercâmbio de *links* e uso das redes sociais, de modo a conseguir os tão pretendidos *backlinks*.

O *Entrevistado E* define a SEO como uma série de ferramentas e de etapas que um *webmaster* percorre para posicionar o seu *site* em um motor de busca como o Google, o Bing ou o Yahoo. As técnicas SEO que usa atualmente para os seus *sites* são submissão a diretórias e raramente *linkbuilding*. Também menciona que as suas preferências por técnicas como: *post* patrocinados⁴¹, *post exchange*⁴² e *article submission*,⁴³ aumentaram. Sublinha finalmente que o conteúdo é o mais relevante de um *site* mesmo tendo em consideração as outras técnicas existentes.

As entrevistas revelaram que as técnicas de submissão a diretórias e *linkbuilding* são as mais usadas pelos entrevistados. No entanto, aparentemente a *linkbuilding* está a ser esquecido, uma vez que dois dos entrevistados mencionaram que raramente o usam. De facto, as notícias mais recentes sobre SEO confirmam que a *linkbuilding* está a perder espaço na otimização de *sites* e que uma nova técnica chamada *linkearning* estaria a surgir [SeoMoz Blog 2012]. O que significa que atualmente os motores de busca estão a dar mais preferência ao conteúdo do que às outras técnicas de posicionamento. Isto não quer dizer que os motores de busca, nunca deram preferência ao conteúdo, no entanto, quando eles atualizam os seus algoritmos é porque estão a pensar em fornecer conteúdo de qualidade aos seus utilizadores [Google Webmaster Blog 2012].

Outro ponto importante nestas entrevistas são as mídias sociais que se têm tornado em ferramentas úteis para obter tráfego e *backlinks*. Os *bookmarkings* sociais, os *blogs* e as redes sociais são um bom exemplo de técnicas SEO correntes que os intervenientes

⁴¹ Um *post* patrocinado é aquele que um *blogger* escreve no seu *blog* sobre os produtos ou serviços de uma empresa a câmbio de um pagamento. Na maioria dos casos, o artigo publicado baseia-se na experiência pessoal do autor com o produto.

⁴² Tratam-se de artigos de *blogs* que contêm *links* dos seus próprios *sites* no texto e que são *linkados* por outros *sites*.

⁴³ Uma técnica usada para melhorar as posições dos *sites* nos motores de busca. Trata-se de escrever um artigo sobre um *site* específico e subseqüentemente submete-lo às diretórias de artigos que existem na *web*.

têm mencionado. Por outro lado, nota-se que os elementos *on-page*, ainda são considerados elementos relevantes como um dos entrevistados afirma: uma boa estrutura de um *site* é suficiente para gerar tráfego e observa a SEO como algo relativo, uma vez que existem outras técnicas de posicionamento. Neste sentido, aparentemente as técnicas SEO tem sido vistas por alguns *webmasters* apenas como uma estratégia de intercâmbio de *links* e não como uma otimização de *sites* desde a sua concepção até depois do seu lançamento.

Como uma medida para evitar algumas penalizações ou declínio do posicionamento nos motores de busca, devido às atualizações constantes realizadas pelos motores de busca, as técnicas SEO como: *post* patrocinados⁴⁴, *post exchange*⁴⁵ e *article submission*⁴⁶, são atualmente as mais valorizadas. Estas técnicas são as novas alternativas de SEO que os *webmasters* estão a empregar e a sua utilização depende das necessidades do *site* a ser otimizado.

Um outro ponto a realçar é a preferência dos entrevistados pelos motores de busca. Todos eles mencionaram que os seus *sites* eram otimizados para o Google, e as técnicas SEO que usavam eram pensando nele. Inclusive um dos intervenientes definiu a SEO como uma otimização de um *site* para o Google. Isto faz-nos refletir que o Google tal como se tinha mostrado nos dados estatísticos fornecidos por ComScore [2012] e StatCounter [2012] continua a ser o líder dos motores de busca na *web*, e que a maioria dos *webdesigners* ou *webmasters* trabalham indirectamente para o Google, uma vez que os *sites* são otimizados para fornecer informação aos seus utilizadores, e que de alguma forma estes também são os utilizadores dos *sites*. Porém, devido às últimas atualizações que o Google tem realizado e ao incremento de *Adwords* nos seus resultados de busca, um dos participantes mencionou que o motor de busca tinha mudado a sua forma de trabalho e, mais do que pensar na qualidade de experiência do utilizador, estavam a pensar em si próprios tendo-se tornado numa ferramenta meramente comercial.

⁴⁴ Um *post* patrocinado é aquele no qual um *blogger* escreve no seu *blog* sobre os produtos ou serviços de uma empresa a troco de um pagamento. Na maioria dos casos, o artigo publicado baseia-se na experiência pessoal do autor com o produto.

⁴⁵ Trata-se de artigos de *blogs* que contem *links* dos seus *sites* no texto e que são *linkados* por outros *sites*.

⁴⁶ Uma técnica usada para melhorar as posições dos *sites* nos motores de busca. Trata-se de escrever um artigo sobre um *site* específico e subseqüentemente submete-lo às diretorias de artigos que existem na *web*.

4.2. Resultados da Avaliação

As técnicas *White Hat* SEO foram avaliados em quatro *sites*. Três deles, conforme foi mencionado nos capítulos anteriores são *sites* com diferentes temáticas que foram solicitados a diferentes donos de *sites* apenas para fins de estudo, e quanto ao último é um *site* de teste criado como parte complementar deste estudo. O processo da avaliação consistiu de três etapas 1) Implementação das técnicas, 2) Monitorização e 3) Relatórios (Ilustração 13).

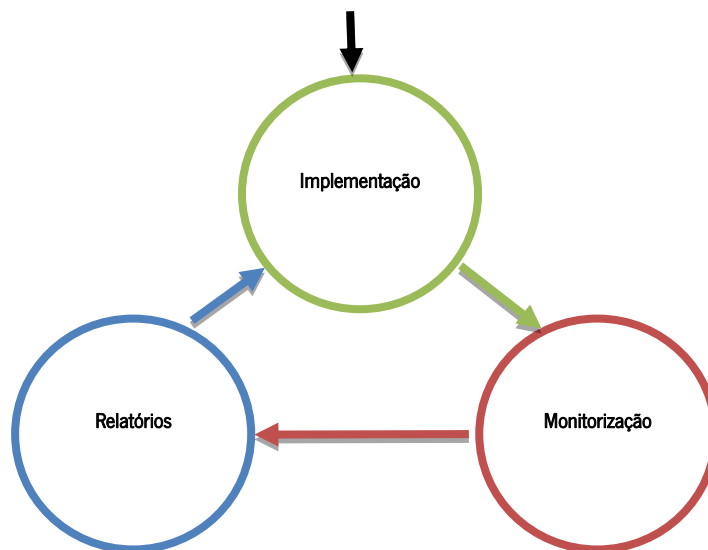


Ilustração 13. Processo de Avaliação das Técnicas SEO

Fonte: Própria

As fases de Implementação e Monitorização foram mencionadas no capítulo anterior. Neste capítulo são apresentados mais detalhes destas fases conjuntamente com os resultados. Posteriormente exibem-se as etapas da avaliação das técnicas SEO através de tabelas e gráficos, de modo a tornar possível o acompanhamento passo a passo todo o processo.

Para os Sites A, B e C

1. Implementação das Técnicas

A implementação das técnicas SEO sobre os *sites* tem sido realizada segundo os relatórios de verificação descritos no capítulo anterior. Os relatórios expõem os elementos de cada *site* com as observações necessárias para começar com a otimização. Da verificação feita nos três *sites* confirma-se que alguns dos elementos que impedem a visibilidade nas SERP são idênticos para cada *site*. Se focarmos nossa atenção para as *tags*, verificamos que nenhum dos *sites* as tem considerado como relevantes, uma vez que usam inadequadamente palavras-chave no *title* ou nenhum *heading* no corpo do texto, ou ainda a duplicação de *meta tags*. A percepção do uso correto das *tags* não irá posicionar os *sites* no topo dos resultados de busca, mas irá ajudar a ser encontrado mais facilmente pelos *crawlers*, para além de ser um meio que fornece informação sobre o tópico da página.

Weideman et al. [2006] dá uma visão mais abrangente sobre esse assunto quando analisa elementos de 144 *sites*. Em 92,3% destes *sites* fez-se uso do *title tag* com termos pouco significativos, em 73,4% colocaram-se menos de 10 termos e em 96,5% as palavras-chave foram escritas em letras minúsculas. O *title tag* em comparação as outras etiquetas HTML é um elemento de elevado valor para a indexação das páginas, pelo que se deve ter cuidado na sua utilização. O uso das palavras-chave em letra minúscula não é um problema de indexação, normalmente estas são escritas dessa maneira para manter um formato harmonioso da página, mas é um problema quando estas são excedidas. Os motores de busca como o Google já não consideram *meta keywords tag* como relevantes porque houve uma época em que os praticantes *Black Hat* SEO fizeram mau uso deste elemento preenchendo-o com uma infinidade de palavras-chave, a fim de posicionar os seus *sites* em pouco tempo.

Um das motivações para estes *sites* serem avaliados foi que a grande maioria dos *webmasters*, constroem *sites* sem levar em consideração esses pequenos detalhes como são as *tags* e depois tentam otimiza-los, ficando frustrados sem saber o porquê de os seus *sites* não conseguirem tráfego, ou porque

estes não foram indexados, ou até porque foram penalizados pelos motores de busca. Este tipo de responsáveis de *sites* não se tem apercebido que a otimização começa antes do *site* ser criado. Pode parecer um pouco falacioso pensar em otimização antes de criar um *site*, mas na realidade é assim como funcionam as coisas. Por isso, antes de lançar um *site* é preciso refletir no que se pretende com o *site*, nas suas funcionalidades, e obviamente, o nome do domínio que é também determinante para o posicionamento.

Este processo SEO foi apresentado, por se crer ser o mais razoável, uma vez que providencia critérios básicos de como começar a otimização de um *site* desde a pesquisa de palavras-chave até a apresentação de relatórios. A tabela 4 a seguir indica os tempos da implementação das técnicas SEO sobre os três *sites* analisados:

	Mês 1	Mês 2	Mês 3	Mês 4
Site A	<i>Tagging, URLs curtas. Criação de sitemap XML</i>	<i>Link internos</i>	Submissão a diretorias	<i>Bookmarking social</i>
Site B	<i>Tagging, Criação de sitemap XML</i>	<i>Link internos</i>	Submissão a diretorias	<i>Bookmarking social</i>
Site C	<i>Tagging,</i>	<i>Link internos</i>	Submissão a diretorias	<i>Bookmarking social</i>
Mês 1: Junho, Mês 2: Julho, Mês 3: Agosto, Mês 4: Setembro				

Tabela 4. Período da Implementação das Técnicas SEO nos *Sites* Analisados.

2. Monitorização das Técnicas

A avaliação das técnicas sobre os *sites* A, B e C focou-se basicamente no número de visitas, número de páginas rastreadas, páginas indexadas, cliques, impressões e consultas, por serem as métricas base fornecidas pelas ferramentas *analytics*. Nos parágrafos a seguir descreve-se a monitorização realizada sobre as técnicas implementadas nos *sites*. As descrições giram à volta dos tempos indicados na tabela 4.

As técnicas *Tagging*, URLs curtas e Criação de *sitemap XML* no primeiro caso provocaram 32 consultas, 700 impressões e 16 cliques no *site A*, tal como é mostrado na ilustração 14. Com a estruturação de *links* internos, no mês a seguir, o número de consultas aumentou para 34, o número de impressões para 1300 e o número de cliques para 22. A submissão a diretórias no terceiro mês incrementou o número de consultas e o número de impressões para 46 e 1300 respectivamente, porém o número de cliques diminuiu para 16. No último mês a *bookmarking* social incrementou favoravelmente o número de consultas para 85, o número de impressões para 2000 e o número de cliques para 35. A ilustração a seguir nos dá uma ideia da evolução do *site A*, uma vez implementadas as técnicas.

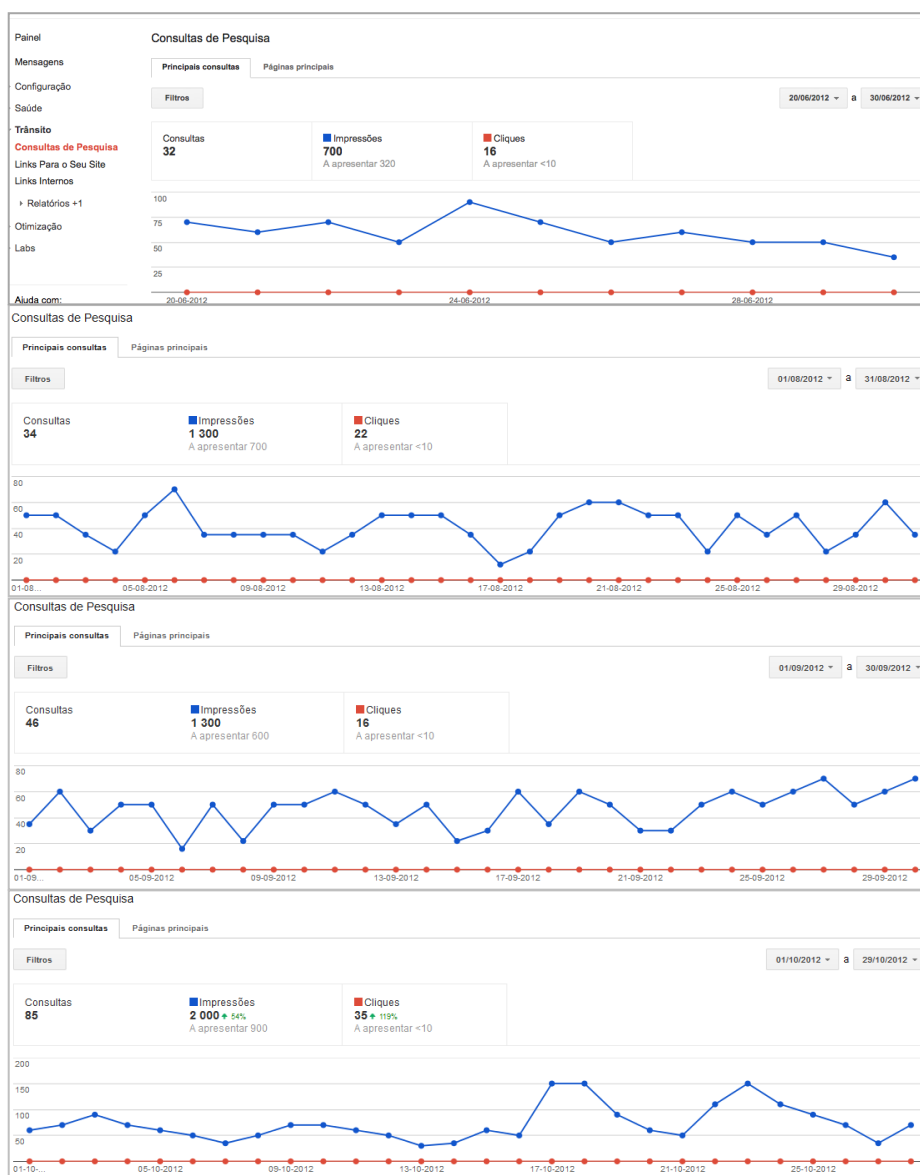


Ilustração 14. Evolução do *Site A* quanto o Tráfego no *Google Webmaster Tool*.

No caso do *site B*, as técnicas *Tagging* e Criação de *sitemap XML* provocaram 19 consultas, 500 impressões e 50 cliques. A estruturação de *links* internos no mês a seguir originou 11 consultas, nota-se que sofreu um abaixamento, o número de impressões manteve-se em 500 e o número de cliques também sofreu um abaixamento para 35. A submissão a diretorias no terceiro mês incrementou o número de consultas e o número de impressões para 19 e 600 respetivamente, mas diminui o número de cliques para 30. No último mês a *bookmarking* social incrementou o número de consultas para 21, e o número de impressões e o número de cliques preservou-se em 600 e 30 respetivamente. A ilustração 15 apresenta a tendência do *site B* durante o processo da implementação das técnicas.

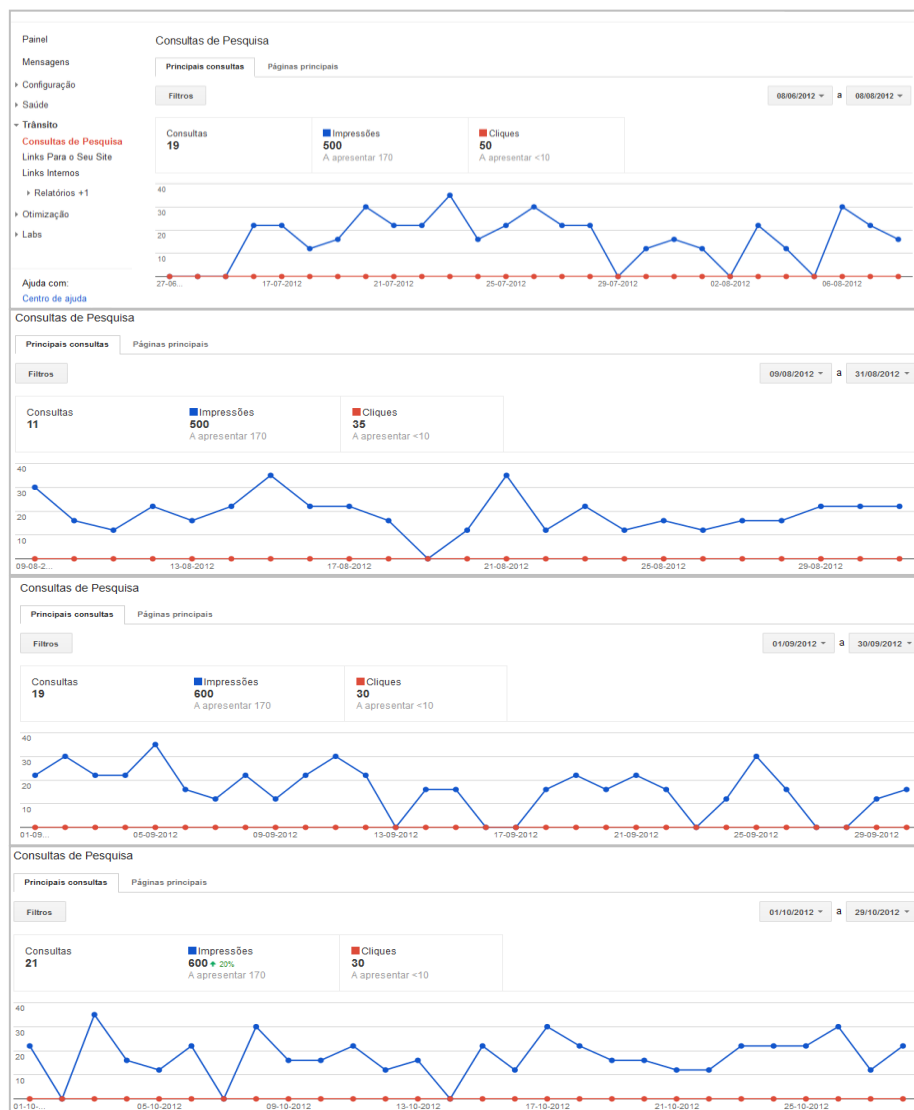


Ilustração 15. Evolução do *Site B* quanto o Tráfego no *Google Webmaster Tool*.

No caso do *site C*, a técnica *Tagging* originou 4 consultas, 150 impressões e menos de 10 cliques. No mês a seguir, a estruturação de *links* gerou um tráfego de 4 consultas, 200 impressões e 22 cliques. Com a técnica submissão a diretorias no terceiro mês, o número de consultas incrementou-se para 7, o número de impressões diminuiu para 170 ao igual que o número de cliques para 16. No último mês a *bookmarking* social originou uma diminuição no número de consultas em 5, o número de impressões manteve-se em 170 e o número de cliques aumentou para 22. A seguir exibe-se a tendência do *site C* durante o processo da implementação das técnicas SEO (Ilustração 16).

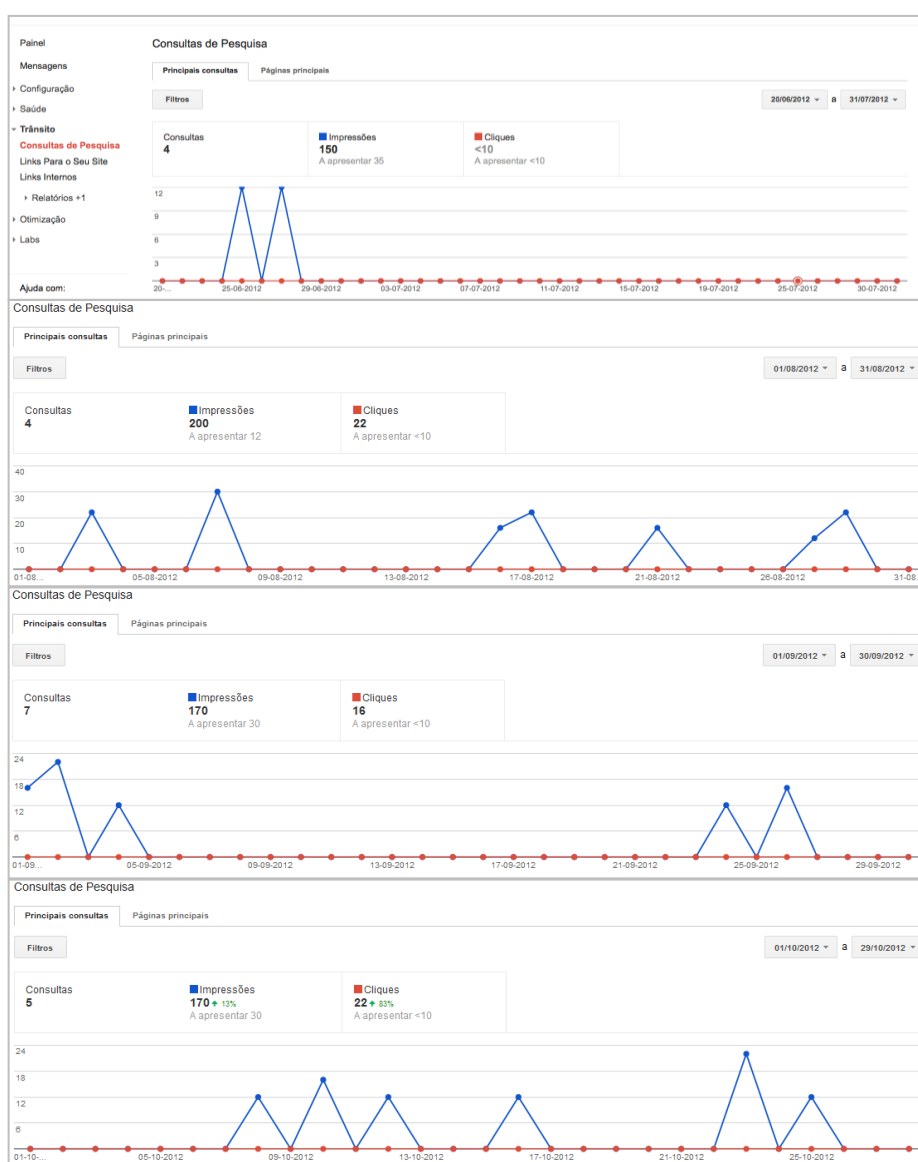


Ilustração 16. Evolução do *Site C* quanto o Tráfego no *Google Webmaster Tool*.

Os gráficos apresentados pelo *Google Webmaster Tool* revelam claramente como o *site A*, B e C têm evoluído com a introdução das técnicas SEO. Nas ilustrações apresentadas observa-se que o número de consultas em alguns casos foi incrementando mês a mês, bem como as impressões e o número de cliques, mas em outros foi ao contrário. Um *background* semelhante observa-se com Bing Webmaster Tool (Ilustração 17). A ilustração a seguir revela a evolução do *site A* nestes últimos dois meses depois da implementação da técnica *bookmarking social*.

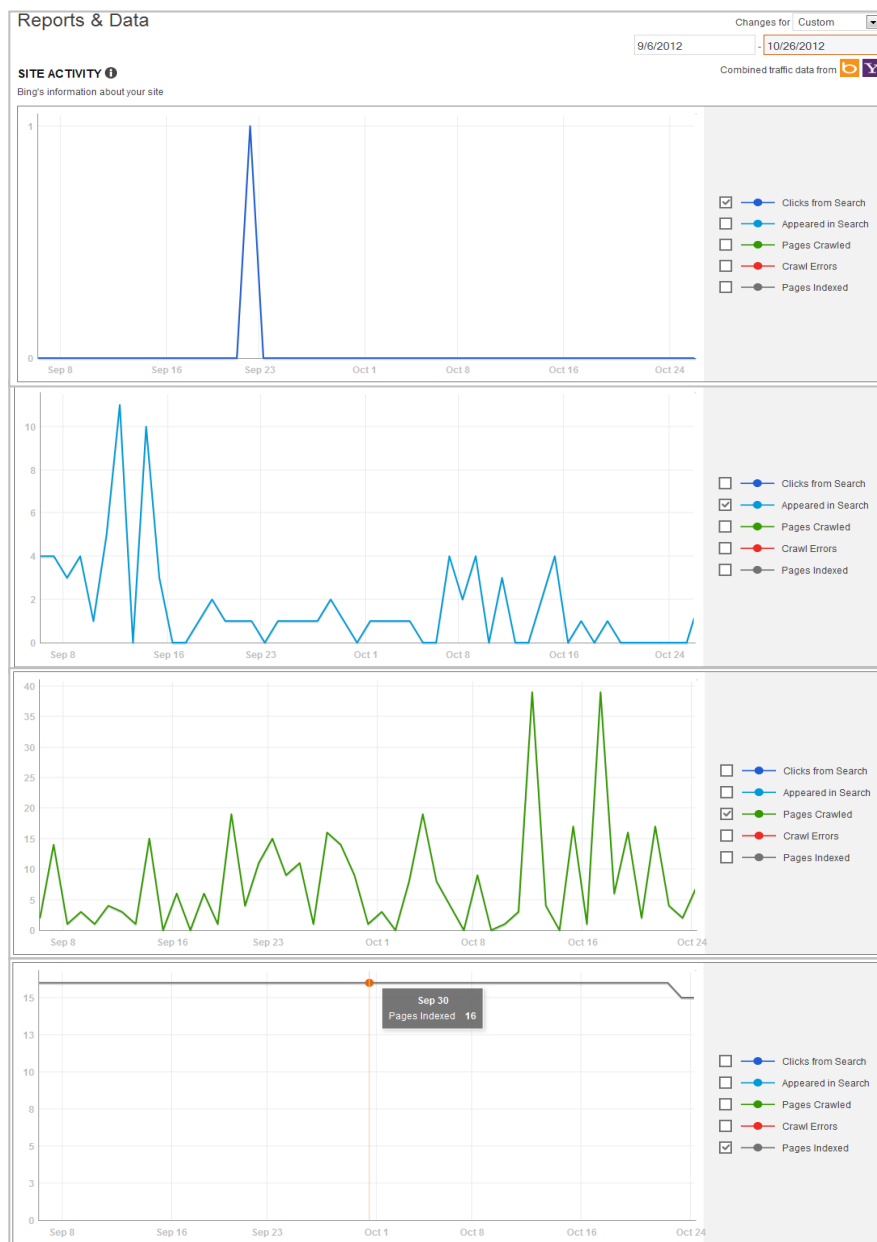


Ilustração 17. Evolução do *Site A* quanto o Tráfego no *Bing Webmaster Tool*.

Vista a ilustração, confirma-se que no mês de Setembro, o *site A* revelou um aumento no número de cliques e no número de impressões. Quanto ao número de páginas rastreadas, os resultados foram aceitáveis, uma vez que houve um aumento no mês de Outubro. Relativamente ao número de páginas indexadas também se manteve de maneira regular, tendo um declínio apenas no fim de Outubro. Os gráficos fornecidos pelo *Bing Webmaster Tool* são semelhantes para os *sites B e C*, pelo que se crê que não há necessidade de reverter informação dos gráficos subsequentes.

3. Relatórios

Nesta etapa apresentam-se os relatórios dos *sites* com os dados estatísticos sobre o tráfego que têm gerado a implementação das técnicas SEO recorrendo-se ao *Google Analytics*. Subsequentemente, exibem-se os gráficos dos relatórios (Ilustração 18, 19 e 20):

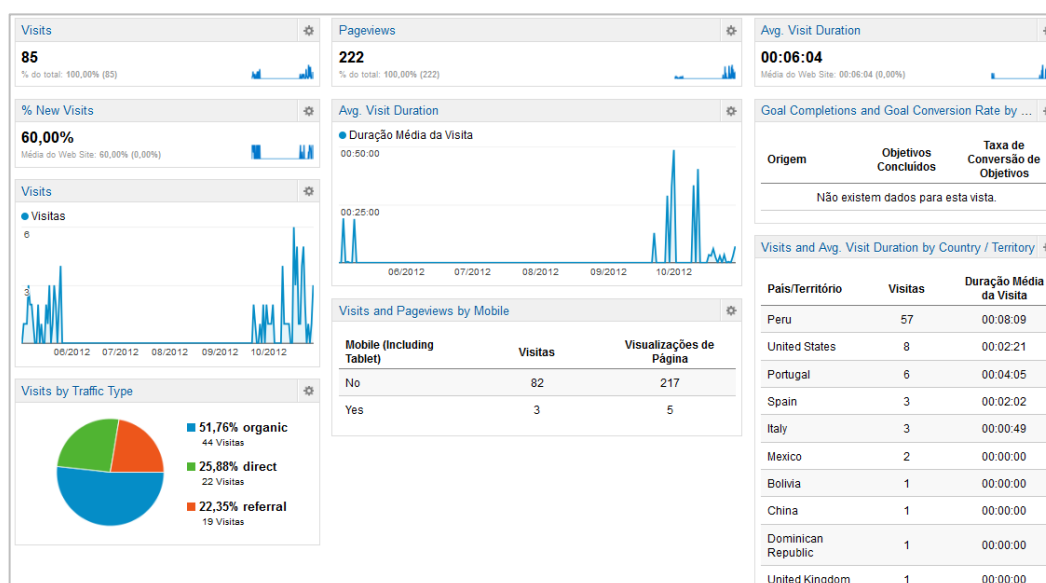


Ilustração 18. Informação Estatística do Tráfego Gerado no *Site A*

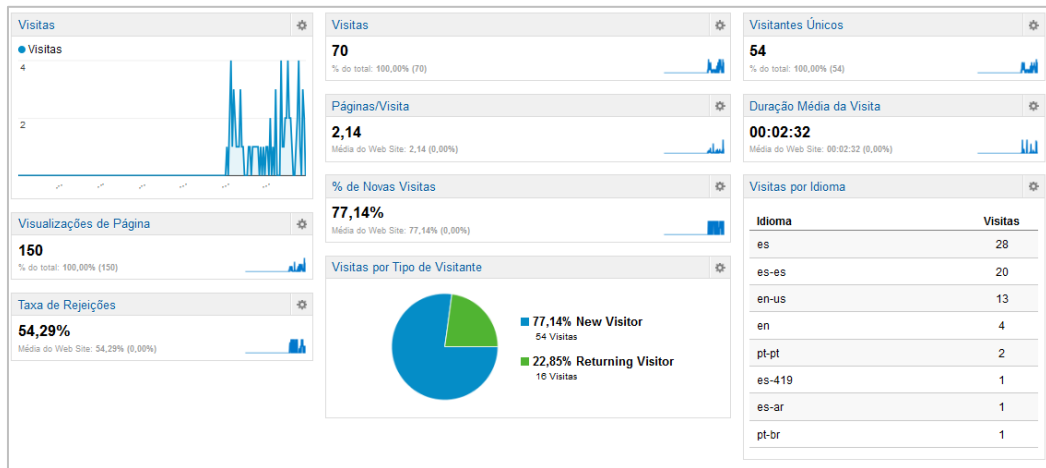


Ilustração 19. Informação Estatística do Tráfego Gerado no *Site B*.



Ilustração 20. Informação Estatística do Tráfego Gerado no *Site C*.

Em seguida é apresentada uma tabela (Tabela 5) com os dados da avaliação das técnicas SEO nos *sites A, B e C*.

		Site A				Site B				Site C			
Técnicas SEO	Métricas	Tagging, URLs curtas, Sitemap XML	Link internos	Submissão a diretórias	Bookmarking social	Tagging, Sitemap XML	Link internos	Submissão a diretórias	Bookmarking social	Tagging	Link internos	Submissão a diretórias	Bookmarking social
		Google	Consultas	32	34	46	85	19	11	19	21	4	4
Impressões	700		1300	1300	2000	500	500	600	600	150	200	170	170
Cliques	16		22	16	35	50	35	30	30	<10	22	16	22
Número de visitas	32		0	8	45	0	0	30	41	0	0	57	79
Páginas visualizadas	37		0	35	150	0	0	55	96	0	0	22	140
Yahoo e Bing	Impressões	0	0	54	30	0	0	9	8	0	0	37	6
	Cliques	0	0	1	0	0	0	0	0	0	0	0	0
	Páginas rastreadas	0	0	127	233	0	0	62	102	0	0	68	96
	Páginas indexadas	0	0	16	15	0	0	8	9	0	0	13	10

Tabela 5. Resultados das Técnicas SEO em Base as Métricas Obtidas da Análise dos Sites A, B e C.

Os dados apresentados na tabela 5 foram extraídos das ferramentas *Google Webmaster Tool*, *Google Analytics* e *Bing Webmaster Tool*. Cada ferramenta fornece um conjunto de métricas que facilita a análise do tráfego e performance dos sites. Desde que a Yahoo e Microsoft fizeram um acordo, o *Bing Webmaster Tool* providencia dados estatísticos do tráfego a partir de ambos motores de busca (Bing Community, 2011), portanto as métricas apresentadas na tabela são dirigidas tanto para o Yahoo como para o Bing.

Para o Site de Teste

1. Implementação

No que respeita às técnicas SEO implementadas no *site* de teste, durante o processo tentou-se seguir rigorosamente sugestões de alguns autores e as orientações dos motores de busca, especificamente do Google por ser o motor de busca com maior número de utilizadores e *sites* indexados [ComScore 2012, StatCounter 2012]. Inicialmente pensou-se no nome do domínio, para este fim recorreu-se à pesquisa de palavras-chave, tal como sugere Jerkovic [2009]. A pesquisa de palavras-chave é um assunto que deve ser levado em consideração, porque é uma forma de compreender o que os utilizadores estão a procurar, mesmo antes de pensar sobre a recolha do conteúdo [Ascensão 2011]. Alguns dos métodos para pesquisar palavras-chave é através de conversas com os utilizadores, entrevistas ou grupos de discussão, afirma McVittie [2012]. Neste sentido, a ferramenta que facilitou à pesquisa de palavras-chave foi a Google Keywords Sugestion Tool.

Uma vez adquirido o nome de domínio www.erasmusexperiences.com prosseguiu-se com o *mock-up* do *site*. A criação do *mock-up* ajudou a ter uma perspetiva geral do que iria ser a estrutura do *site* de teste. Miller [2000] contribui para esta criação com alguns dos fatores que lista no seu artigo *The Website Quality Challenge* sobre como estruturar um *site* de qualidade. Finalmente o *site* de teste foi desenvolvido com base no *mock-up* realizado, sendo assim a página de entrada (ilustração 16) e a página utilizador (ilustração 17) como segue:

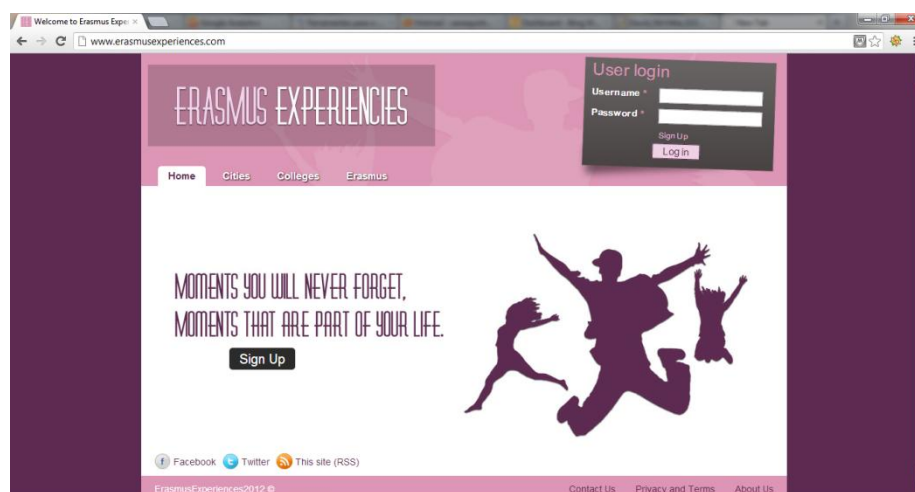


Ilustração 21. Página de Entrada do Site de Teste.



Ilustração 22. Página Utilizador do *Site* de Teste.

O desenvolvimento implicou a utilização das técnicas, linguagem de estilo, *tagging*, conteúdo próprio e *bookmarkings* sociais. Os tempos da introdução das técnicas no *site* de teste (Tabela 6) são indicados a seguir:

	Semana 1	Semana 2
<i>Site</i> de Teste	Linguagem de estilo, <i>Tagging</i>	Conteúdo próprio, <i>Bookmarking social</i>
Semana 1- Semana 2: As semanas correspondem à segunda e terceira semanas de Setembro		

Tabela 6. Período da Implementação das Técnicas SEO no *Site* de Teste

Importa mencionar que o nome do domínio do *site* de teste foi adquirido no mês de Abril e no mês a seguir colocou-se apenas um esboço do *site*. Nos meses a seguir até Agosto criaram-se algumas funcionalidades, sendo finalmente publicado na terceira semana de Setembro.

2. Monitorização

A avaliação das técnicas SEO no *site* de teste focou-se nas seguintes métricas: Número de visitas, número de páginas rastreadas, páginas indexadas, cliques, impressões e consultas. Segundo a Tabela 6, na primeira semana fez-se uso da técnica *linguagem de estilo*. Esta técnica aplicou-se ao *design* do *site*, uma vez que o uso das CSS permite separar a formatação das páginas *web* do seu conteúdo, fazendo com que estas se tornem ágeis. Na semana a seguir, implementou-se a técnica *tagging* para que o *site* fosse rastreado facilmente pelos *crawlers*. Na terceira semana aplicou-se a técnica de conteúdo próprio criando uma conta no *site* para introduzir informação sobre a experiência Erasmus da autora. E finalmente, aplicou-se a técnica dos *bookmarkings* sociais a cada página do *site* a fim de obter mais tráfego e assim posicioná-lo nos motores de busca já mencionados. Uma vez implementadas as técnicas obteve-se informação do número de visitas, páginas vistas por visita, duração das visitas, entre outros, tal como se mostra na ilustração a seguir:

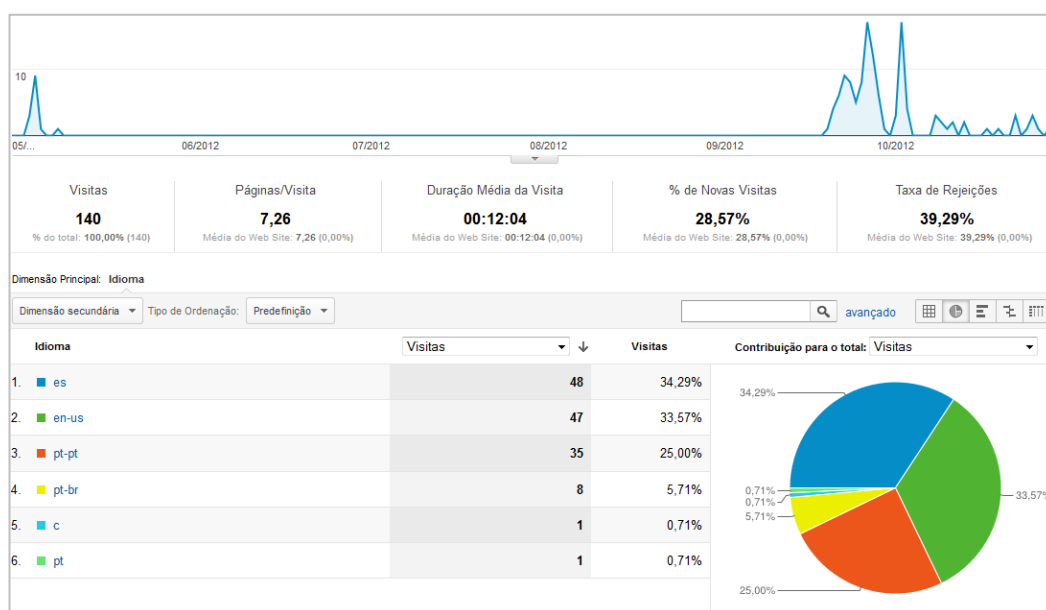


Ilustração 23. Evolução do *Site* de Teste desde a sua Criação.

Inicialmente, como pode visualizado na ilustração 19, o *site* obteve uma percentagem aceitável de visitas no início de Maio, depois de ter sido publicado o esboço. Nos meses posteriores, analisada a situação e com o propósito de não perder o tráfego, publicou-se uma mensagem no qual se afirmava que o *site* estava

sob construção. Finalmente o *site* foi publicado em meados de Setembro, gerando tráfego poucos dias depois de ser lançado. As técnicas SEO foram utilizadas durante a criação do *site* e depois de ele ser lançado.

3. Relatórios

Em seguida apresenta-se uma tabela (Tabela 7) com dados da avaliação das técnicas SEO no *site* de teste.

		Site de Teste	
		Técnicas SEO	Mettricas
		Linguagem de Estilo, URLs Curtas, <i>Tagging</i>	Conteúdo Proprio, <i>Bookmarking</i> , social
Google	Consultas de busca	0	1
	Impressões	<10	22
	Cliques	<10	<10
	Número de visitas	20	57
	Páginas visualizadas	136	432
Yahoo e Bing	Impressões	2	76
	Cliques	0	0
	Páginas rastreadas	10	109
	Páginas indexadas	3	10

Tabela 7. Resultados das Técnicas SEO no *Site* de Teste

Como no caso da tabela 5, os dados estatísticos apresentados nesta tabela foram extraídos das ferramentas *Google Webmaster Tool*, *Google Analytics* e *Bing Webmaster Tool*. Uma vez realizado o processo de implementação e monitorização dos tempos mencionados, obteve-se um relatório final da tendência atual do *site* (Ilustração 24).

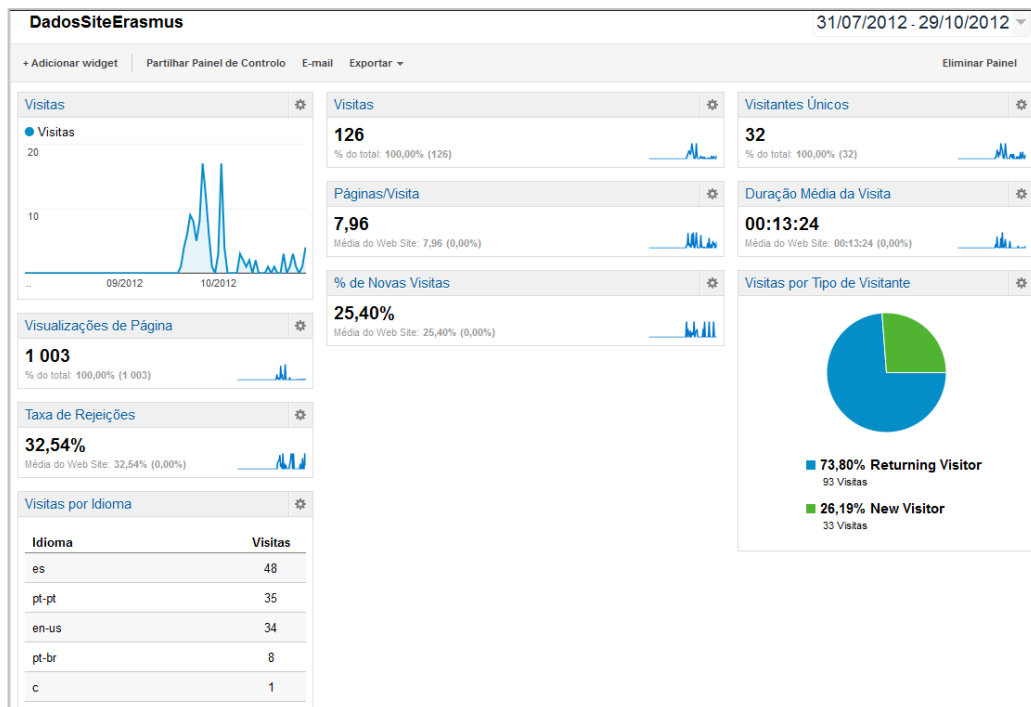


Ilustração 24. Informação Estatística do Tráfego Gerado no *Site* de Teste.

O *site* atualmente tem conseguido 126 visitas, 1003 visualizações da página com uma duração média da visita de 13 minutos e 24 segundos, provenientes da língua espanhola, portuguesa, inglesa e chinesa. Visto estes dados estatísticos, fez-se uma consulta do *site* nos três diferentes motores de busca para verificar a posição na que se encontrava. Para este fim utilizou-se a palavra-chave “Erasmus Experiences”. Numa primeira instância, a palavra-chave consultada no motor de busca Google posicionou ao *site* no lugar dezoito dos 3 860 000 resultados, tal como é mostrada na Ilustração 25. A mesma palavra-chave consultada no Bing situou ao *site* na posição dois dos 16 800 000 resultados de busca (Ilustração 26). Finalmente dos 7 370 000 resultados de busca obtidos do Yahoo depois de ter consultado a palavra-chave, verificou-se que o *site* situava-se no segundo lugar.

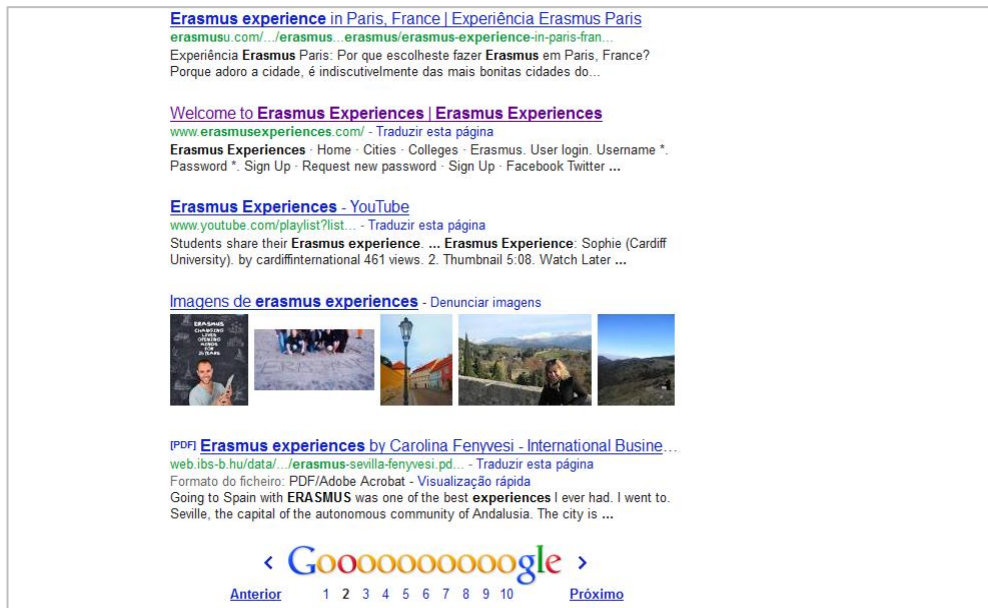


Ilustração 25. Posição do Site de Teste no Google.

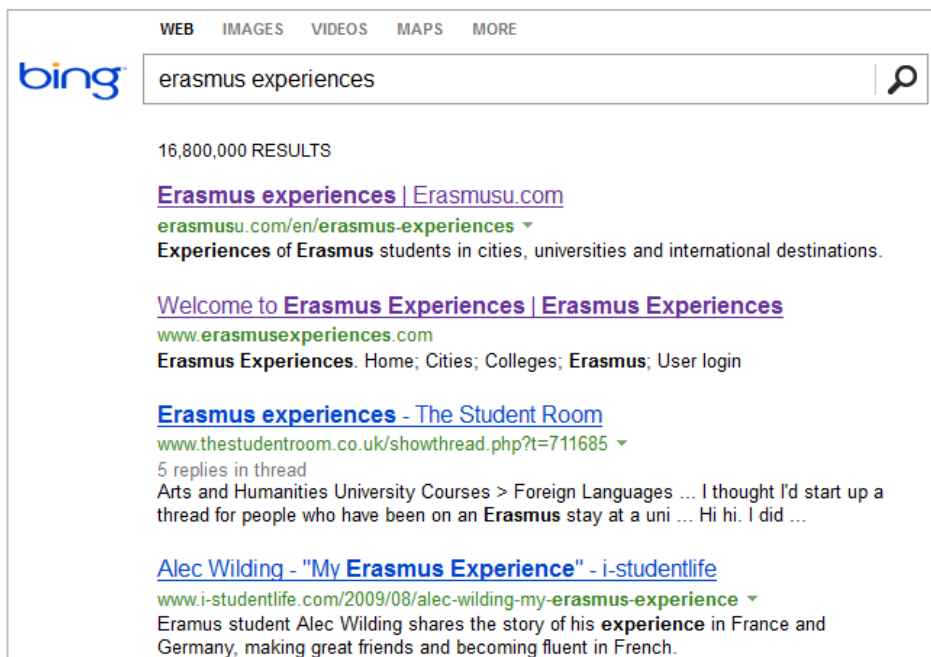


Ilustração 26. Posição do Site de Teste no Bing.

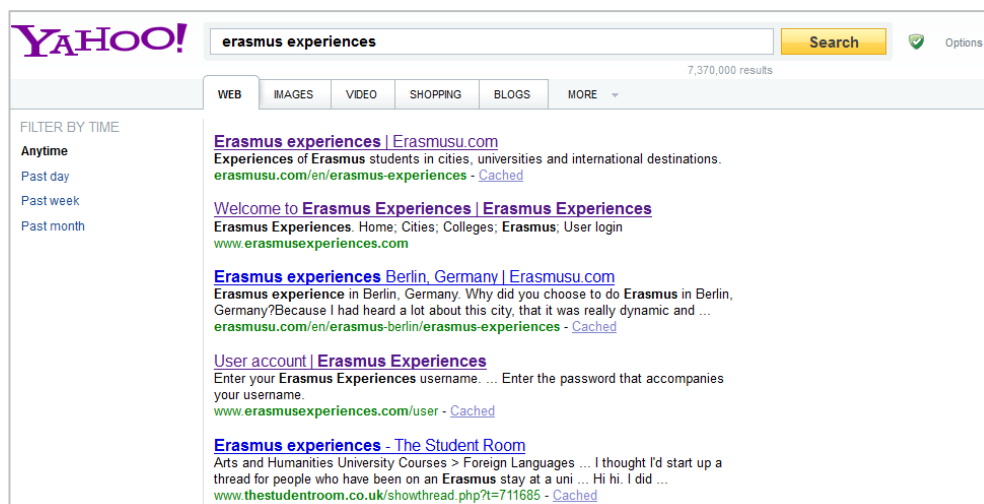


Ilustração 27. Posição do Site de Teste no Yahoo.

4.3. Comparação das Técnicas SEO

Uma vez obtida os resultados das técnicas implementadas e monitorizadas os criterios considerados para avaliar a eficacia das técnicas foram:

1. Incremento de tráfego
2. Facilidade de rastreamento pelos motores de busca
3. Posicionamento nas primeiras páginas dos motores de busca

Segundo a tabela 5 as técnicas *taggings*, URL curtas e a criação de *sitemaps* XML em conjunto providenciam um tráfego favorável e mantem o código simples, enquanto os *links* internos não se mostram como relevantes na geração de tráfego. A submissão a diretorias mantém-se como uma das técnicas SEO relevantes uma vez que os dados fornecidos mostram um incremento na maioria das métricas apresentadas. O *bookmarking* social nota-se que tem conquistado a área, como se pode observar é a técnica que provoca um aumento maior de tráfego comparativamente com as restantes.

Segundo a tabela 7 as técnicas: Linguagem de estilo, URLs curtas e *tagging* aplicam-se em conjunto para uma primeira avaliação e as outras técnicas: conteúdo próprio e *bookmarking* social empregam-se em conjunto para uma segunda avaliação. Os dados das métricas apresentadas variam conforme as técnicas são implementadas,

pelo que se pode afirmar que os conjuntos apresentados são os mais adequados para gerar tráfego nos *sites*.

Analisados os resultados das duas tabelas apresentadas e aos critérios apresentados inicialmente se conclui que as técnicas SEO mais eficazes são: *Linguagem de estilo*, uma vez que agiliza o *site*, separando a formatação do conteúdo; *tagging* que ajuda a indexar ao *site*, através das descrições fornecidas nos *tags* HTML; a URLs curtas que permite aos *crawlers* rastrear os *links* mais rapidamente sem causar erros; a *bookmarking* social, é um elemento relevante com que se pode obter *backlinks* e portanto incrementar tráfego. Pode-se dizer que é uma das técnicas SEO mais recentes que está a ser utilizado na maioria dos *sites*, particularmente nos blogs; e por ultimo a técnica do conteúdo próprio. Diz-se sempre que o conteúdo num *site* é “o rei” por ser o elemento fundamental que providencia informação aos utilizadores. É por esta razão que os motores de busca tentam de providenciar informação de qualidade aos utilizadores, alterando constantemente os seus algoritmos para reduzir o *spam*. A seguir expõe-se a tabela 8 com as técnicas SEO mais eficazes.

Categorias	Técnicas <i>SEO</i> mais Eficazes
Conteúdo	<ul style="list-style-type: none">• Conteúdo Próprio
Código	<ul style="list-style-type: none">• Linguagem de Estilo• <i>Tagging</i>• URLs Curtas
Ligações	<ul style="list-style-type: none">• <i>Bookmarking</i> Social

Tabela 8. Técnicas SEO Correntes mais Eficazes.

Uma vez feita a comparação das tabelas 5 e 7, a tabela 8 apresenta um resultado final deste estudo com as técnicas SEO que se consideram mais eficazes. Esta tabela nos da uma ideia das técnicas SEO base que correntemente podem beneficiar aos *sites* na *web*.

Capítulo 5 - Discussão

5.1. Avaliação das Técnicas

As técnicas SEO têm surgido como um meio de otimização para os motores de busca e como tal têm-se tornado em boas práticas para o desenvolvimento e promoção dos *sites* nos motores de busca.

Os resultados apresentados no capítulo anterior dão uma percepção de que as técnicas SEO se podem aplicar de forma individual ou em conjunto nos *sites*. Tudo depende das necessidades do *site* e o critério que cada *webmaster* coloca no *site*. Quando alguém vai adquirindo experiência na área, alcança o conhecimento e a capacidade de deduzir quais são as técnicas mais benéficas para o *site*. Igualmente, existem ferramentas SEO que ajudam a identificar as técnicas que o *site* precisa, uma vez que estas verificam o estado inicial do *site*, a sua performance e o tráfego que está a adquirir.

Entre as técnicas SEO apresentadas como eficazes observa-se que a técnica de criar conteúdo próprio é a mais relevante. Falar de conteúdo próprio não implica necessariamente copiar conteúdo de vários *sites* e colocá-lo como próprio, mais do que isso é fornecer informação que as pessoas estão a procurar. As vezes, parece impossível criar conteúdo novo entre os milhares de *sites* que existem na *web*, mas não é assim, cada *site* é único seja no seu *design*, conteúdo, ou código. Portanto, pensar em conteúdo próprio consiste em utilizar os recursos que estão a nossa disposição para depois partilhar esse conteúdo com pessoas de interesses comuns, ou mesmo obter conteúdo novo das próprias pessoas. Hoje em dia com o avanço da tecnologia web 2.0 isto é mais simples, os utilizadores têm-se tornado nos próprios editores de conteúdo, uma vez que interagem continuamente, dando opiniões, sugestões e comentários, seja em *blogs*, fóruns ou nas redes sociais. Uma boa ideia é aproveitar esta interação para gerar conteúdo único. É claro que muitos outros *sites* só procuram benefícios próprios e são demasiado otimizados, não tendo em consideração a experiência dos utilizadores. Neste caso, são os motores de busca que decidem se eventualmente penalizam, ou não, este tipo de *sites*.

Quanto à linguagem de estilo, as *taggins* e as URLs curtas, são uma boa maneira de estruturar o código fonte e agilizar a apresentação do *site*. Como é bem conhecido, os utilizadores quando estão a procurar informação na *web* são bastante impacientes, não gostam de esperar ou permanecer muito tempo num *site*, se este não oferece o que precisam. Pelo que, a boa estruturação destas técnicas irá manter o *site* ligeiro, acelerando a sua apresentação em poucos milissegundos. Além de que ajudará os *crawlers* indexar as páginas de maneira mais simples e rápida.

No que respeita aos *bookmarkings* sociais, são o que correntemente se está a usar para obter tráfego e *backlinks*. É uma técnica simples de usar nos *sites* que permite aos utilizadores partilhar as suas preferências de conteúdo nas redes sociais. Pode-se dizer que é uma técnica chave para obter tráfego nos *sites* enquanto existem as redes sociais.

Criar um *website* de qualidade que atrai *backlinks* de maneira natural é uma forma de acrescentar valor ao *site*, uma vez que os algoritmos dos motores de busca como o Google medem a importância ou popularidade dos *sites* através do número de *backlinks*. O foco pode estar nestes desejados *backlinks*, mas no fim, os algoritmos dos motores de busca são os que definem que *sites* e porque irão a ser classificados nas SERP, sendo feitas atualizações de entre um a dois meses [Search Engine Land 2012b]. O motivo das atualizações constantes é reduzir o número de *sites spam* nos resultados de busca, neste sentido as técnicas SEO atuam como o método confiável para criar *sites*.

5.2. Resultados nos Motores de Busca

De acordo com Potts [2007] existem quatro formas de as pessoas procurarem conteúdo na *web* 1) digitando diretamente um URL, 2) clicando em um *link* de referência sobre outro *site*, 3) selecionando o *site* através da página de resultados de um motor de busca, e, 4) utilizando mídias sociais e *sites* de redes sociais. Cada um destes itens representa um meio de comunicação entre as pessoas e a *web* com o objetivo de querer obter informação relevante. Se concentrarmos nossa atenção no item 3, verificamos que é a consulta mais ampla e imparcial, uma vez que, os resultados são classificados por relevância através do uso de algoritmos complexos, que até hoje ninguém tem certeza de como é que eles operam. No entanto, é o meio mais usado pelas

peessoas para obter informação relevante. Neste sentido os motores de busca tornam-se numa ferramenta de grande importância.

Os motores de busca sempre tiveram o objetivo de prestar um bom serviço aos seus utilizadores, devolvendo-lhes conteúdo útil nas suas pesquisas. Porém, com o surgimento das novas tecnologias, a quantidade inumerável de *sites* na *web* e o número de empresas dispostas a investir em publicidade, provocaram com que eles se tornassem ferramentas mercantis ao contrário de serem apenas ferramentas focadas nas necessidades de busca do utilizador. Um resultado disto é a quantidade de publicidade que se encontra nos resultados de busca em forma de *links* a partir das consultas efetuadas pelos utilizadores, tentando passar despercebidos entre *links* que não tem nada a ver com a publicidade. Estes tipos de resultados nos motores de busca distinguem-se como resultados orgânicos e resultados pagos. Os resultados orgânicos ou naturais surgem da utilização das técnicas SEO [McDonald e Zayat 2009] enquanto os resultados pago surgem da publicidade anunciada sobre os motores de busca, normalmente referida como SEM [Work 2008]. A ilustração 4 apresentada no capítulo 2 está demonstrada esta diferença.

Por outro lado, está comprovado que nem todos os utilizadores sabem da existência dos dois tipos de resultados que os motores de busca fornecem. Um estudo da Pew Internet apresentado por Mordkovich e Mordkovich [2005] indicou que o 62% das pessoas não percebem que há dois tipos de resultados nas SERP, e isto só pode indicar três coisas: os utilizadores focam-se mais na informação que precisam e não observam detalhes para além de seus olhos, o *design* do *site* dos motores de busca não faz evidência da existência deste tipo de resultados, ou as pessoas não têm nenhum conhecimento sobre motores de busca. Ao refletir sobre o assunto, assume-se que há pessoas que precisam de ser instruídas na obtenção de resultados, de modo que obtenham resultados genuínos e precisos. Uma proposta interessante é a de Ascensão [2011] que propõe a pesquisa avançada como parte de ensino escolar obrigatório para obter resultados mais precisos. Propostas como esta tentam de ajudar na obtenção de resultados com informação de qualidade, adaptados às necessidades de busca do utilizador.

No que diz respeito ao comportamento dos utilizadores relativamente aos resultados, tem sido demonstrado que os utilizadores raramente olham para além da

terceira página dos resultados de busca e que a maioria deles fazem clique apenas na primeira página. O estudo efetuado por [Malaga 2008] confirma que 62% dos utilizadores de motores de busca clicam apenas nas três primeiras páginas de resultados e que só 10% vai além da terceira página de resultados. Outro estudo revela que 73% dos utilizadores dos motores de busca não olham para além da primeira página [Schimmel, 2010]. Estes indicadores confirmam a real importância para os *webmasters* e *marketeers* quererem posicionar os seus *sites* nos primeiros lugares das SERP. Normalmente a listagem das páginas de resultados abrange dez *links*. Aparecer no topo das dez primeiras, com a enorme quantidade de *sites* existentes na *web*, é todo um desafio. No entanto esta listagem, no último ano, tem sofrido algumas alterações fazendo com que os motores de busca apresentem apenas sete *links* por página. O Google é um dos motores que já apresenta uma listagem como esta na sua primeira página de resultados [Search Engine Land 2012].

5.3. Visibilidade

Obter visibilidade nas SERP dos motores de busca não é tão simples como parece, é preciso tempo, determinação e muita paciência, porque não é um trabalho de resultados imediatos mas sim um trabalho de longo prazo sendo que quando se atingem resultados é bastante satisfatório. Estatísticas como a ITU (International Telecommunications Union) demonstram que um terço da população mundial está conectada a Internet, ou afirmações como de Schooner [2010] que declara que 90% de utilizadores da *Internet* suecos fazem uso de motores de busca ao menos uma vez por semana, fazem-nos refletir que o meio *online* possui um enorme potencial para atrair as pessoas. Portanto, a visibilidade através do uso de motores de busca é um aspeto muito importante dentro do *Marketing Online* uma vez que possibilita a captação de potenciais clientes para o negócio.

Uma vez aprofundado o conhecimento sobre o mundo do *Marketing Online* e SEO, observou-se que existe uma infinidade de fatores que impedem os *sites* de conseguir visibilidade nos motores de busca. Autores como Schooner [2010], Thurow [2007] e Weideman et al. [2006] fazem referência aos fatores que impedem a visibilidade:

- Uso de *Java Script* na estrutura de navegação. Atualmente a maioria dos motores de busca não permitem *links* dentro do código *Java Script*, ou simplesmente põem um limite para eles rastrearem.
- Uso de *frames*.
- Páginas com uso excessivo de palavras-chave.
- Excessivos parâmetros dinâmicos na URL.
- *Backlinks* ligados a *sites* de *spam*.
- Repetição desproporcional do mesmo texto âncora em grandes quantidades a partir de *backlinks*.
- Ligações excessivas de *sites* hospedados no mesmo endereço IP de classe C.
- Não ter incluído *Meta Tags* nas páginas.
- Ter incluído *flash* no *design*.
- Gráficos de grande volume.

Quer sejam elementos externos ou elementos internos de um *site*, estes afetam a forma como ele é classificado em qualquer motor de busca. Uma revisão de cada um deles deveria ser tida em consideração, de modo que a visibilidade de um *site* nos motores de busca não fique tão longe do previsto.

Até este ponto têm-se mencionado aspetos puramente técnicos, focados na satisfação dos motores de busca, mas não se tem tido em consideração que a visibilidade de um *site* depende mais dos utilizadores do que dos motores de busca. Um dos principais erros que os *webmasters*, responsáveis de *sites* ou *marketeers* fazem é, focarem-se mais em questões técnicas do que na experiência dos utilizadores, quando o ponto da viragem está focado sobre eles. Responder aos utilizadores, entender os seus problemas, os seus desejos e as suas necessidades é uma parte importante do *marketing* nos motores de busca. Alguns comentários sobre este aspeto podem-se encontrar no livro “The findability Formula: the easy, non-technical approach to search engine” de [Lutze 2009], que ainda identifica palavras-chave e termos de pesquisa para os *sites* serem encontrados de modo mais personalizado.

5.3. Google e SEO

Alguns responsáveis de *sites* perguntam se realmente o Google está, ou não, a favor do SEO uma vez que está constantemente a atualizar os seus algoritmos e, de uma maneira ou outra, a afetar uma percentagem mínima dos *sites* que se consideram de qualidade. Mas, como podem perder visibilidade os *sites* de qualidade nos resultados de busca? Se o Google está sempre a publicar que os *sites* têm de ser de qualidade, que têm de respeitar as diretrizes, que têm de evitar o uso de técnicas maliciosas e até tem criado um *blog* direcionado para os *webmasters* onde está publicado um guia para iniciantes SEO. Esta perspetiva do Google parece contraditória quando na sua listagem de resultados mostra *sites* com pouca qualidade, no entanto, o que pretende é equilibrar os resultados de busca de modo que este providencie informação útil aos utilizadores. Consequentemente, Google afirma no seu *blog* que as técnicas *white hat* SEO melhoram a usabilidade de um *site*, ajudam a criar um bom conteúdo e fazem com que o *site* funcione com rapidez, o que é bom para utilizadores e motores de busca [Google Webmaster Blog 2012].

No *blog* de *webmasters* do Google pode-se encontrar muitos comentários a favor e contra as atualizações do Google. Alguns *webmasters* afirmam que a ação do Google é positiva, uma vez que alguns *sites* só preferem copiar o esforço que puseram nos seus conteúdos, enquanto outros afirmam que são as piores atualizações que puderam ver. Uma vez percebido que os *webmasters* não chegam a um consenso, pode-se ver que o Google tenta fazer o seu melhor para providenciar conteúdo ótimo aos utilizadores, apesar de no último ano ter-se focado mais nas *adwords*. É obvio que o Google depois de ter praticamente abrangido quase todo o mercado da Internet e ter-se tornado numa empresa sólida, está à procura de mais benefícios próprios e quer restringir os *adsenses* apenas aos *sites* que considerem merecedores.

Capítulo 6 - Conclusões

6.1. Conclusões do Estudo

Segundo os resultados apresentados no capítulo anterior pode-se perceber que o *site* criado obteve maior tráfego em pouco tempo em comparação dos outros *sites*. Portanto, as técnicas SEO deveriam ser usadas desde o início da criação de um *site* e não depois de modo que se possa conseguir visibilidade em pouco tempo.

SEO, mais do que um conjunto de técnicas, são boas práticas que podem existir para otimizar um *site*. Uma ideia errada que se tem criado nestes últimos anos é de que SEO tem de ver apenas com conseguir *backlinks*. Sabe-se que os *backlinks* são um dos elementos mais relevantes para demonstrar a importância de um *site*, mas com todas as mudanças que os motores de busca realizam, não é recomendado focar-se só neste elemento, quando se tem outros elementos nos que também se deve trabalhar. O nome de domínio, a idade do *site*, a estruturação dos *tags*, as URLs, os *meta tags* e o conteúdo também fazem parte do trabalho SEO.

O trajeto típico da maioria dos responsáveis de *sites* é 1) comprar um domínio, sem fazer pesquisas de palavras-chave, 2) criar o *site* sem ter em consideração o estilo apropriado, 3) submeter o *site* em alguns motores de busca e diretorias e finalmente 4) esperar que as pessoas o descubram. Depois de algumas semanas ou meses apercebe-se que o *site* não tem sido encontrado pelo internautas. Pelo que no fim, optam por otimizar o *site* na tentativa de obter visibilidade nos motores de busca.

O que não se tem notado nesse percurso é que a otimização começa desde o nome do domínio de um *site* e não se pode cair num erro de quer otimizar um *site* depois de este ter sido criado. É por isso que existe uma infinidade de *sites* que não se podem posicionar facilmente nos motores de busca, e outra quantidade de responsáveis de *sites* que se perguntam porque o seu *site* não têm obtido visibilidade ou porque este foi banido pelos motores de busca. Portanto, uma das primeiras coisas que deve ser tida em consideração antes de criar um *site* é a pesquisa de palavras-chave.

Por outro lado, segundo as estatísticas apresentadas por ComScore [2012] e StatCounter [2012] e as entrevistas realizadas, o Google tem sido colocado como o líder dos motores de busca. Tal significa que as técnicas SEO são aplicadas mais em função das diretrizes do Google do que dos outros motores de busca.

Quanto às críticas de que as técnicas SEO vão desaparecer pelas constantes mudanças realizadas pelos motores de busca, crê-se que estas irão sempre existir, uma vez que os motores de busca acabam por se basear nestas técnicas para apresentar *sites* com informação de qualidade aos utilizadores. Portanto a hipótese de que podem desaparecer fica muito longe da realidade. O que vai acontecer é que estas vão evoluir, sendo um bom exemplo disto o pedido de introdução de dados estruturados nos *websites* pela Google, como uma forma de compreender melhor o seus conteúdos.

O mais importante de uma atividade de otimização dos motores de busca é perceber frequentemente o que os utilizadores fazem quando estão num *site*. Neste sentido, as ferramentas *analytics* fornecem inúmeros dados estatísticos que permitem analisar o comportamento dos utilizadores, mas muitas vezes os responsáveis de *sites* não sabem como agir com estes dados. O que comumente eles fazem é gerar tráfego nos *sites* cada vez mais sem saber o que fazer depois. Incrementar tráfego nos *sites* não é um erro, este é principal propósito de SEO, mas se vai-se omitir a experiência do utilizador não faz muito sentido manter o *site*, uma vez que os utilizadores são o ponto crucial de que os *sites* existam na *web*. Por isso é importante focar-se nas necessidades dos utilizadores, através de seguimentos contínuos no *site* para aperceber o que lhe mantém no *site* até torna-lo em clientes. A este tipo de ações chama-se otimização da taxa de conversão que permite tornar aos visitantes em clientes, os *sites* comerciais dedicados a vendas de produtos, serviços entre outros focam-se continuamente neste assunto. Monitorizar o estado de um *site* não é um trabalho fascinante mas a recompensa no longo prazo vale a pena.

Segundo o estudo realizado, também é possível perceber que os *sites* que usam as técnicas SEO não podem ficar em pé de igualdade perante os motores de busca, uma vez que cada *site* está focado num tópico diferente e fornece diversos serviços. Alguns *sites* são pessoais, comerciais ou são simplesmente informativos. O uso destas técnicas vai depender muito da capacidade dos responsáveis dos *sites* para obter a visibilidade nos motores de busca.

Quanto ao posicionamento nos motores de busca, segundo as tabelas e as imagens apresentadas no *site* de teste no capítulo 4, está comprovado que a classificação dos *sites* varia. Um *site* pode-se posicionar em primeiro lugar no Bing, em segundo lugar no Yahoo, e pode ficar na trigésima posição no Google. Isto é a prova de que os algoritmos de indexação variam nos vários motores de busca existentes.

Conclui-se com a expectativa de se ter contribuído para o enriquecimento dos conhecimentos na área das tecnologias *Web*.

6.1. Limitações

O estudo focou-se na avaliação das técnicas *White Hat* SEO e não tanto nas técnicas *Black Hat* SEO por estas serem banidas pelos motores de busca. A revisão do estado de arte ajudou a dar uma percepção geral das técnicas.

Os *sites* foram analisados através de ferramentas estatísticas gratuitas, neste caso o *Google Analytics* e de ferramentas de suporte para *webmasters*, o *Google Webmaster Tools* e o *Bing Webmaster Tools*. Outras ferramentas como o *Seo Quake*, *Link Checker* de W3C e *SEOMoz Pro*, que não tem sido mencionadas na revisão de literatura também foram usadas, de modo a complementar o trabalho de otimização e avaliação dos *sites*.

Os nomes dos três *sites* analisados não foram mencionados na documentação por serem *websites* de pessoas externas ao estudo e porque foram providenciadas apenas com fins de estudo.

6.2. Contribuições

Uma das principais contribuições deste trabalho é fornecer informação relevante aos responsáveis de *sites* sobre o uso das técnicas SEO, como é que estas evoluem quando são implementadas de maneira correta, e os benefícios que trazem quando são utilizadas desde o início da construção dos *sites*, sem necessidade de recorrer a técnicas antiéticas. Um exemplo do correto uso das técnicas foi demonstrado no *site* de teste.

Outro contributo é a identificação das técnicas eficazes SEO, se usadas, permitirão uma indexação mais eficiente por parte dos motores de busca.

A avaliação das técnicas SEO efetuada auxilia, também, na obtenção de *sites* mais amigáveis para os utilizadores (usabilidade), pelo que o trabalho contribui tanto para os utilizadores dos motores de busca como para os visitantes dos *sites*.

6.4. Recomendações e Trabalhos Futuros

Para as pessoas interessadas na área recomenda-se estar atento às atualizações que surgem dia a dia, uma vez que SEO é bastante amplo e está a mudar constantemente com as alterações que fazem os motores de busca.

Os motores de busca estão constantemente a mudar as suas políticas de posicionamento. Como trabalho futuro sugere-se validar as técnicas SEO noutras *sites* fora dos já apresentados e focar o estudo das técnicas SEO no âmbito da *web* semântica. O Google pretende neste momento incluir-se na problemática da *Web Semântica*, fornecendo nas opções de *Google Webmaster Tools*, recomendações aos responsáveis de *sites*, para que estes passem a conter dados mais estruturados, de modo que o seu conteúdo possa ser envolvido ao nível da interoperabilidade de semântica na *web*. Portanto, crê-se que as técnicas SEO aplicadas num *site* semântico influenciará de maneira distinta o posicionamento nos motores de busca.

6.5. Considerações Finais

Para os leitores perceberem o tema de investigação, os capítulos apresentados descreveram com minuciosidade o trabalho efetuado durante o estudo, pelo que se espera seja de total compreensão para as futuras investigações.

Nota-se que o estudo das técnicas SEO neste trabalho foi orientado para as pessoas ou empresas que pretendem conseguir tráfego nos seus *sites*, ou que pretendem criar um *site* como base de um negócio.

Referências

- Arnold, D. (2010). *PPC Marketing for Beginners*. Spruce Goose Enterprises.
- Ascensão, C. P. (2011). *Google Marketing - A mais poderosa arma para antigir os seus clientes* (Sílabo). Manuel Robalo.
- Bing. (2012). Guidelines for Successful Indexing. Retrieved August 30, 2012, from <http://onlinehelp.microsoft.com/en-us/bing/hh204434.aspx>
- Bing Community. (2011). Announcing the integration of Yahoo traffic data into Bing Webmaster Tools reports. *Webmaster Center Blog*. Retrieved August 31, 2012, from http://www.bing.com/community/site_blogs/b/webmaster/archive/2011/08/16/announcing-the-integration-of-yahoo-traffic-data-into-bing-webmaster-tools-reports.aspx
- Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Computer networks and ISDN systems*, 30(1-7), 107–117.
- Bröcker, J., & Van Ahee, G. (2008). *Semantics & Search Engine Optimisation*.
- Burgemeister, S. (2003). *Market analysis*. GRIN Verlag.
- Chaffey, D. (2008). *eMarketing eXcellence: Planning and optimizing your digital marketing*. Routledge.
- Chaters, B. (2011). *Mastering Search Analytics: Measuring SEO, SEM and Site Search*. O'Reilly Media.
- Chu, H., & Rosenthal, M. (1996). Search engines for the World Wide Web: A comparative study and evaluation methodology. In *Proceedings of the Annual Meeting-American Society for Information Science* (Vol. 33, pp. 127–135). Retrieved from http://cui.unige.ch/tcs/cours/algoweb/2002/articles/art_habashi_arash.pdf
- ComScore. (2012). comScore Releases June 2012 U.S. Search Engine Rankings. Retrieved August 30, 2012, from http://www.comscore.com/Press_Events/Press_Releases/2012/6/comScore_Releases_June_2012_U.S._Search_Engine_Rankings
- Costa, A., Rodrigues, E., Pinto, F., Macedo, J., & Nicolau, M. J. (1995). *Internet-Guia Prático do Cibernauta. Campo das Letras*.
- Curran, K. (2004). Tips for achieving high positioning in the results pages of the major search engines. *Information Technology Journal*, 3(2), 202–205.
- Cutroni, J. (2010). *Google analytics*. O'Reilly Media, Inc.

DAA. (2008). Web Analytics Association Web Analytics Definitions. Retrieved from http://c.yimcdn.com/sites/www.digitalanalyticsassociation.org/resource/resmgr/PDF_standards/WebAnalyticsDefinitions.pdf

David, M. (2011). *Wordpress 3.0 Search Engine Optimization*. Packt Publishing Ltd.

Emtage, A., & Deutsch, P. (1992). Archie: An electronic directory service for the internet. In *Proceedings of the Winter 1992 Usenix Conference* (pp. 93–110).

Enaceanu, A. (2008). SEO Techniques for Business Websites. *Journal of Information Systems & Operations Management*, 2(1), 114–121.

Enge, E., Spencer, S., Fishkin, R., & Stricchiola, J. C. (2009). *The art of SEO*. O'Reilly Media, Inc.

Evans, M. P. (2007). Analysing Google rankings through search engine optimization data. *Internet Research*, 17(1), 21–37. doi:10.1108/10662240710730470

Fernandes, E. M. (2001). Grelha de repertório. University of Minho. Retrieved from <http://repositorium.sdum.uminho.pt/handle/1822/4210>

Fister, J. M. (2007). Correlation analysis of on-page attributes and search engine rankings. University of Cincinnati, Engineering : Computer Science & Engineering.

Fleiss, W. (2007). *The Evolution of Search Engine Optimization*. Bennett Kuhn Varner, Inc.

Gandour, A., & Regolini, A. (2011). Web site search engine optimization: a case study of Fragformet. *Library Hi Tech News*, 28(6), 6–13.

Gay, R., Charlesworth, A., & Esen, R. (2007). *Online marketing: a customer-led approach*. Oxford University Press, USA.

George, D. (2005). *The ABC of Seo*. Lulu. Com.

Google. (2012). General Guidelines. Retrieved August 30, 2012, from <http://support.google.com/webmasters/bin/topic.py?hl=en&topic=2370419&parent=1724125&ctx=topic>

Google Analytics. (2012). Google Analytics Features. Retrieved August 30, 2012, from <http://www.google.com/analytics/features/index.html>

Google Webmaster. (2012). Creating Sitemaps. Retrieved August 30, 2012, from <http://support.google.com/webmasters/bin/answer.py?hl=en&answer=183668>

Google Webmaster Blog. (2012). Another step to reward high-quality sites. Official Google Webmaster Central Blog. Retrieved September 30, 2012, from <http://googlewebmastercentral.blogspot.pt/2012/04/another-step-to-reward-high-quality.html>

- Gulli, A., & Signorini, A. (2005). The indexable web is more than 11.5 billion pages. In *Special Interest Tracks and Posters of the 14th International Conference on World Wide Web* (pp. 902–903).
- ICT. (2011). The World in 2011: ICT Facts and Figures. Retrieved from <http://www.itu.int/ITU-D/ict/facts/2011/material/ICTFactsFigures2011.pdf>
- IWS. (2012, August 31). Internet World Users by Language. *Top Ten Internet Languages - World Internet Statistics*. Retrieved from <http://www.internetworldstats.com/stats7.htm>
- Jerkovic, J. I. (2009). *SEO warrior*. O'Reilly Media, Inc.
- John, J. P., Yu, F., Xie, Y., Krishnamurthy, A., & Abadi, M. (2011). deSEO: Combating search-result poisoning. *Proceedings of the 20th USENIX Security*.
- Jones, K. B. (2008). *Search Engine Optimization: Your visual blueprint for effective Internet marketing*. Wiley Publishing, Inc.
- Kent, P. (2008). *Search engine optimization for dummies*. Wiley Publishing, Inc.
- Kobayashi, M., & Takeda, K. (2000). Information retrieval on the web. *ACM Computing Surveys (CSUR)*, 32(2), 144–173.
- Kothari, C. R. (2009). *Research methodology: methods and techniques*. New Age International.
- Kremer, J. M. (1980). A técnica do incidente crítico. *Revista da Escola de Biblioteconomia da UFMG, Belo Horizonte*, 9(21), 165–76.
- Langville, A. N., & Meyer, C. D. (2003). Fiddling with pagerank. *Center for Research in Scientific Computation (CRSC), North Carolina State University, Report CRSC-TR03-34*. Retrieved from http://meyer.math.ncsu.edu/Meyer/PS_Files/FiddlingPageRank.pdf
- Ledford, J. (2009). *SEO: search engine optimization bible*. Wiley Publishing, Inc.
- Ledford, J., Teixeira, J., & Tyler, M. E. (2010). *Google analytics*. Wiley Publishing, Inc.
- Lewandowski, D. (2005). Web searching, search engines and Information Retrieval. *Information Services and Use*, 25(3), 137–147.
- Lieb, R. (2009). *The truth about search engine optimization*. Que Publishing Inc.
- Lieberam-Schmidt, S. (2010). *Analyzing and Influencing Search Engine Results: Business and Technology Impacts on Web Information Retrieval*. Gabler Verlag.

- Lutze, H. F. (2009). *The findability formula: The easy, non-technical approach to search engine marketing*. John Wiley & Sons Inc.
- MacDonald, M. (2011). *Creating a Website: The Missing Manual: The Missing Manual*. O'Reilly Media, Inc.
- Malaga, R. A. (2008). Worst practices in search engine optimization. *Communications of the ACM*, 51(12), 147–150.
- Mann, C., & Stewart, F. (2000). *Internet communication and qualitative research: A handbook for researching online*. Sage Publications Limited.
- Marie, K. (2012). *Mafia: Killer Marketing for Indie Authors*. eBookIt.com.
- Maxwell, J. A. (2004). *Qualitative research design: An interactive approach*. Sage Publications, Incorporated.
- McDonald, T., & Zayat, F. (2009). The Internet as a marketing medium: An introduction to Search Engine Optimization. Mikkeli University of Applied Sciences.
- McVittie, D. (2012). SEO Tools: Looking at SEO Specialists as a User Class. University of Guelph.
- Metaxas, P. T., & DeStefano, J. (2005). Web spam, propaganda and trust. In *The 1st International Workshop on Adversarial Information Retrieval on the Web (AIRWeb)*.
- Miletsky, J. I. (2009). *Principles of Internet Marketing: New Tools and Methods for Web Developers*. Course Technology Ptr.
- Miller, E. (2000). The Website Quality Challenge. *Software Research, Inc.*
- Miller, M. (2012). *B2B Digital Marketing: Using the Web to Market Directly to Businesses*. Que Publishing.
- Mitchell, M. (2012). *SEO & PPC: Better Together*. Pearson Education, Inc.
- Mordkovich, B., & Mordkovich, E. (2005). *Pay-per-click search engine marketing handbook: low cost strategies to attracting new customers using Google, Yahoo & other search engines*. Lulu. com.
- Muljadi, E. P. (2011). *Digital Marketing Handbook*. Paul Muljadi.
- Nykiel, R. (2007). *Handbook of marketing research methodologies for hospitality and tourism*. Routledge, Inc.
- Peterson, E. T. (2004). *Web analytics demystified: a marketer's guide to understanding how your web site affects your business*. Celilo Group Media.

- Pineda, J. M., Paraskevas, A., Frew, A. J., & others. (2004). Web Analytics: A Missed Opportunity for Hotel Online Strategy. *Information and Communication Technologies in Tourism*, ed. AJ Frew, Springer, Vienna, 623–633.
- Pinkerton, B. (2000). *Webcrawler: Finding what people want*. University of Washington.
- Pizam, A. (2010). *International Encyclopedia of Hospitality Management*. Routledge Inc.
- Potts, K. (2007). *Web Design and Marketing Solutions for Business Websites*. Friends of ED.
- Princz, M. (2007). Search engine ranking. *7th International Conference on Applied Informatics*, 2, 417–422.
- Rugg, G., & Petre, M. (2006). *A gentle guide to research methods*. Open University Press.
- Schimmel, R. (2010). Keyword Suggestion for Search Engine Marketing. University of Twente.
- Schooner, P. (2010). Model for Auditing Search Engine Optimization for E-business. University of Linköping, Department of Science and Technology.
- Search Engine Land. (2012a). Google Panda Update 20 Released, 2.4% Of English Queries Impacted. Retrieved October 30, 2012, from <http://searchengineland.com/google-panda-update-20-released-2-4-of-english-queries-impacted-135291>
- Search Engine Land. (2012b). 7 Is The New 10? Google Showing Fewer Results & More From Same Domain. Retrieved August 30, 2012, from <http://searchengineland.com/7-new-10-google-showing-fewer-results-131006>
- SeoMoz. (2012). SEO Best Practices. Retrieved August 30, 2012, from <http://www.seomoz.org/learn-seo/internal-link>
- SeoMoz Blog. (2008). The Role of Outbound Links. *The Daily SEO Blog*. Retrieved August 31, 2012, from <http://www.seomoz.org/blog/the-role-of-outbound-links>
- SeoMoz Blog. (2012). The Death of Link Building and the Rebirth of Link Earning - Whiteboard Friday. *The Daily SEO Blog*. Retrieved September 30, 2012, from <http://www.seomoz.org/blog/the-death-of-link-building-and-the-rebirth-of-link-earning-whiteboard-friday>
- Spindler, S. (2010). *Online Marketing: How to Increase International Sales with Search Engine Optimisation*. GRIN Verlag.

- StatCounter. (2012). Top 5 Search engines in the United States from Jul 2011 to Jul 2012. Retrieved August 30, 2012, from http://gs.statcounter.com/#search_engine-US-monthly-201107-201207
- Straub, D., Boudreau, M., & Gefen, D. (2004). ISWorld Section on Quantitative, Positivist Research. *Quantitative Research in Information Systems*. Retrieved October 15, 2012, from <http://dstraub.cis.gsu.edu:88/quant>
- Sullivan, D. (2003). How search engines rank Web pages. *Search Engine Watch*, 31.
- Svore, K. M., Wu, Q., Burges, C. J. C., & Raman, A. (2007). Improving web spam classification using rank-time features. In *Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web* (pp. 9–16).
- Thurrow, S. (2007). *Search engine visibility*. Peachpit Press.
- Tuten, T. L. (2008). *Advertising 2.0: Social media marketing in a Web 2.0 world*. Praeger Publishers.
- Vaishnavi, V., & Kuechler, W. (2004). Design research in information systems. Retrieved from <http://www.citeulike.org/group/4795/article/6505471>
- Véronis, J. (2006). A comparative study of six search engines. URL: <http://www.up.univmrs.fr/veronis/pdf/2006-comparative-study.pdf>. Retrieved from http://usuarios.multimania.es/webbuscadores/2006_comparative_study.pdf
- Weideman, M., Chambers, R., & Bannister, F. (2006). Improving Website Visibility and Information Retrieval of E-Commerce Ventures: a Specification to Please the Crawlers. *Informatics & Design Papers and Reports*. Paper 93. http://dk.cput.ac.za/inf_papers/93
- West, A. G., Hayati, P., Potdar, V., & Lee, I. (2012). Spamming for science: Active measurement in Web 2.0 abuse research. In *WECSR '12: Proceedings of the Third Workshop on Ethics in Computer Security Research*.
- Wohlin, C., Höst, M., & Henningson, K. (2003). Empirical research methods in software engineering. *Empirical Methods and Studies in Software Engineering*, 7–23.
- Work, J. (2008). *Scientific Search Engine Marketing*. Search Engine Marketing.
- Wynne, P. (2012). *Pimp My Site: The DIY Guide to SEO, Search Marketing, Social Media and Online PR*. John Wiley & Sons.
- Yahoo. (2012). Webmaster resources. Retrieved August 30, 2012, from http://help.yahoo.com/kb/index?page=content&y=PROD_SRCH&locale=en_US&id=SLN2248&impressions=true

Anexos

Anexo I – Mensagem Eletrónica de Pedido de Entrevista

Boa Tarde,

O meu nome é Alexandra Quintana e sou investigadora no Departamento de Sistemas de Informação da Escola de Engenharia da Universidade do Minho. Presentemente, participo num estudo sobre "Search Engine Optimization Techniques Evaluation", em que se pretende avaliar as técnicas SEO que são utilizadas hoje em dia para otimizar os sites e assim conseguir maior visibilidade nos motores de busca. O supervisor académico deste estudo é o Prof. Filipe de Sá-Soares (fss@dsi.uminho.pt) do Departamento de Sistemas de Informação.

O objectivo do estudo consiste em identificar e caracterizar as técnicas mais eficazes de SEO com o propósito de criar um equilíbrio entre o que os motores de busca solicitam e o que os sites pretendem, uma vez que o algoritmo dos motores de busca está em constante mudança.

Uma das fases deste projecto de investigação envolve a realização de entrevistas de cerca de 15 minutos a peritos SEO para se ter uma noção mais clara sobre as técnicas consideradas mais eficazes e a opinião dos entrevistados relativamente às diretrizes dos motores de busca.

Neste sentido, gostaria de solicitar a vossa colaboração, e se a resposta for afirmativa saber a vossa disponibilidade para a realização de uma entrevista.

Em momento algum da entrevista se solicitará informação reservada sobre a vossa organização, sendo que as questões se cingirão à recolha da visão e opinião do entrevistado sobre as actuais técnicas de SEO.

Desde já agradeço a vossa atenção.

Na expectativa de uma resposta positiva ao meu pedido de colaboração, apresento os meus cumprimentos.

Alexandra Quintana

Departamento de Sistemas de Informação

Escola de Engenharia

Universidade do Minho

4800-058 GUIMARÃES

Anexo II – Guião da Entrevista

Qual o cargo que exerce dentro da sua empresa?

Pode fazer uma apresentação da sua empresa?

Que compreende por SEO?

Quais as técnicas SEO mais usadas na sua empresa?

Como mede a qualidade de um *site*?

Quais são as métricas que utilizam para medir a qualidade de um *site*?

O que ferramentas utilizam para medir o medir a qualidade de um *site*?

Anexo III – Resumo das Entrevistas

Entrevistado A

Termo SEO: Otimização para motores de busca, aparecer na primeira página dos resultados dos motores de busca através de *linkbuilding*, redes sociais ou outras tecnologias que possam surgir.

Técnicas SEO que usa: *Bookmarkings* sociais, *linkbuilding* e *blogs*.

Ferramentas SEO que usa: Ferramentas de SEOMoz, SeoQuake para medirem a qualidade de uma página; Google Analytics e Google Webmaster Tools para verificar o tráfego e desempenho do *site*.

Técnica SEO mais Eficazes: Conteúdo dinâmico e verificação da otimização *on-page* e *off-page*.

Que achas que vai acontecer com SEO: *“Alguns acham que já deixuo de existir há muito tempo, mas eu acho que ainda continua a existir, só que têm de se adaptar às novas diretrizes dos motores de busca, neste caso o Google que é o mais predominante. Portanto, SEO vai continuar a existir enquanto se adaptar às normas dos motores de busca”.*

Comentários finais: *“SEO tenta jogar com os resultados dos motores de busca mas sem SEO os motores de busca não conseguiram determinar quais são páginas de qualidade, por isso acho que SEO e os motores de busca têm uma relação bastante estreita”*

Entrevistado B

Termo SEO: Otimizar elementos *on-page* e *off-page* de uma página para os motores de busca especialmente para o Google.

Técnicas SEO que usa: Uma boa estrutura de *links* internos, *sitemap*, bom conteúdo, *blogs*, e intercâmbio de *links*.

Ferramentas SEO que usa: Google Analytics para verificar o Tráfego, Google Webmaster Tools para ver como se tem indexado a página, e ferramentas de seomoz para monitorar as palavras-chave.

Técnicas SEO mais Eficazes: tudo o que está contido na otimização *on-page*, *sitemap*, boa estrutura de *links* internos, bom conteúdo e artigos de *blogs*.

Que esperas dos motores de busca: *“Que os motores de busca continuem a melhorar, que os resultados sejam cada vez mais orientados ao conteúdo e que não apresentem só uma listagem de resultados mas um resumo do que precisa. Isto é a nova tendência, e vejo que já é muito utilizado”*

Comentários finais: *“Que se continue a pesquisar as novas tendências SEO e sobretudo apresentar um bom conteúdo ”*

Entrevistado C

Termo SEO: SEO, otimização de *sites* para os motores de busca, algo com que muitas pessoas estão a lidar.

Técnicas SEO que usa: Bom *design*, boa estrutura de código fonte baseando-se em padrões, *tags*, palavras-chave e *links* externos.

Ferramentas SEO que usa: Ferramentas de Google.

Técnica SEO que considera Eficazes: Ter ma boa estrutura do *site*, *backlinks* e conteúdo.

O que parecer sugere-te os motores de busca frente ao SEO: *“como já tinha mencionado tudo está gira à volta do Google, eles sempre estão a criar novas funcionalidades e novos projetos para mostrar resultados cada vez mais personalizados aos utilizadores, neste sentido SEO em si é muito relativo mas serve de todas formas”*

Comentarios finais: *“Sugiro que as pessoas dedicadas a criar sites não se comprometam a fazer vários trabalhos ao mesmo tempo porque só pode acontecer duas coisas: os sites não sejam concluídos no tempo combinado ou que sejam concluídos com pouca qualidade. Isto faz que o trabalho de criar sites seja desvalorizado. E que os clientes nunca têm razão, sempre querem mais do que é necessário, pelo que seria bom instruí-los e dar algumas dicas do que precisam”.*

Entrevistado D

Termo SEO: SEO é o acrónimo de Search Engine Optimization e pode-se referir tanto à otimização de *sites* para conseguir visibilidade no topo dos resultados dos motores de busca, bem como à pessoa responsável de fazer esse trabalho.

Técnicas SEO que usa: Conteúdo relevante, submissão nos diretórios, intercâmbio de enlces e uso das redes sociais.

Ferramentas SEO que usa: Google Analytics, Google Webmaster Tool, Google Keywords Suggestion Tool que sugere palavras-chave para encontrar novos nichos de mercado, algumas ferramentas de SeoMoz, entre outras que facilitam o trabalho do *Webmaster*.

Técnicas SEO que considera Eficazes: Submissão nos diretórios e Intercâmbio de enlces

O que achas que vai acontecer com o SEO: *“Acho que vai evoluir, hoje não é mais como era antes e daqui alguns anos não vai continuar a ser o que é agora. Antes usavam-se alguns métodos de otimização que hoje são penalizadas pelos motores de busca. Por isso acho que no futuro o SEO vai evoluir e não desaparecer, porém vai mudar a sua forma de trabalho”*

Comentários Finais: *“Os resultados não são de curto prazo mas sim a médio prazo. É um trabalho que está em constante mudança, pelo que temos de ter paciência e devemos estar sempre atentos às atualizações mais recentes de SEO. E não esqueçamos que o mais importante neste trabalho é a qualidade de Conteúdo.*

Entrevistado E

Termo SEO: SEO é uma série de ferramentas e de etapas que um Webmaster prossegue para posicionar o seu *site* em um motor de busca quer no Google, quer no Bing ou no Yahoo.

Técnicas SEO que usa: Submissão nos diretórios, *link building*, *post patrocinados*, *post exchange*, *article submission*, e conteúdo.

Ferramentas SEO que usa: ferramentas de SEOmoz, SeoQuake, Majestic SEO, Google analytics e Google Webmasters tools

Técnicas SEO que considera Eficazes: *Post exchange* e *article submission*.

Que esperas dos motores de busca: *“Que o algoritmo que têm esteja segundo as diretrizes que eles mesmos propõem, porque na realidade isto não está acontecer”. Por exemplo o Google está-se a tornar cada vez mais comercial. Com as últimas alterações que fez no seu algoritmo, os resultados ficaram muito díspares para algumas páginas”.*

Que achas que vai acontecer com o SEO: *“Vai continuar a evoluir e a melhorar os resultados dos motores de busca. Portanto o SEO continuará a existir, tendo em consideração outros fatores de otimização mas sempre focado no conteúdo.”*

Comentários Finais: *“O conteúdo é o mais importante nas páginas e definitivamente se uma pessoa quer fazer SEO tem de ler muito, ou seja, andar a par com as atualizações mais recentes de todos os motores de busca”.*

Apêndice

Apêndice I – Diretrizes dos Motores de Busca mais Conhecidos

A. Google

O Google tem dividido as suas diretrizes nestes três itens [Google 2012].

1. Diretrizes de *design* e conteúdo

- Construa um *site* com uma hierarquia e *links* de texto⁴⁷ claros. Cada página deve ser acessível por pelo menos um *link* de texto estático.
- Apresente aos utilizadores um *sitemap* com *links* que conduzam aos tópicos mais importantes do seu *site*. Se o *sitemap* tiver muitos *links*, convém dividi-lo em várias páginas.
- Empregue um número razoável de *links* por página.
- Crie um *site* útil, com muitas informações e páginas que descrevam de forma clara e precisa o seu conteúdo.
- Pense nas palavras que os utilizadores digitariam para encontrar as suas páginas e verifique se elas aparecem no seu *site*.
- Tente usar texto, em vez de imagens, para exibir nomes, conteúdo ou *links* importantes. O rastreador do Google não reconhece texto contido em imagens. Se for necessário usar imagens para conteúdo de texto, avalie a possibilidade de usar o atributo "ALT" para incluir algumas palavras de texto descritivo.
- Verifique se os elementos `<title>` e os atributos ALT são descritivos e precisos.
- Verifique se há *links* quebrados e corrija o código HTML.
- Se decidir usar páginas dinâmicas (ou seja, a URL contém um carácter "?"), seja ciente de que nem todos os *spiders* de motores de busca rastreiam as páginas dinâmicas e estáticas. Sugere-se manter os parâmetros curtos e a quantidade desses parâmetros pequena.

⁴⁷Conhecido também como um hipertexto ou ligação, que redireciona ao utilizador de uma seção ou documento a outra dentro de uma página através de um clique.

2. Diretrizes Técnicas

- Use um *browser* de texto como o Lynx⁴⁸ para examinar o seu site, pois muitos *spiders* de mecanismos de busca vêm o site do mesmo modo que o Lynx. Se recursos especiais como *JavaScript*, *cookies*, IDs de sessão, *frames*, DHTML (*Dynamic HTML*) ou *Flash* lhe permitirem ver todo o site em um *browser* de texto, o mais provável é que os *spiders* dos motores de busca terão dificuldade em rastrear o seu site.
- Permita que os robôs de busca rastreiem o seu *site* sem que os IDs de sessão ou os argumentos acompanhem suas trajetórias no site. Estas técnicas são úteis para monitorar um determinado comportamento do utilizador, mas o padrão de acesso dos robôs é completamente diferente. O uso dessas técnicas pode gerar uma indexação incompleta, já que os robôs provavelmente não conseguirão eliminar os URLs que têm aspecto diferente, mas que apontam para a mesma página.
- Verifique se o servidor *web* suporta o cabeçalho "*If-Modified-Since HTTP*". Este recurso permite que o servidor *web* informe ao Google se houve alteração no conteúdo desde o último rastreamento do *site*. A utilização deste recurso economiza largura de banda e a sobrecarga.
- Use o ficheiro robots.txt no servidor *web*. Este ficheiro informa aos *Crawlers* quais as diretorias que podem ou não ser rastreadas. Verifique se ele está disponível no *site* de modo que não bloqueie acidentalmente o rastreador do *Googlebot*. Pode testar o seu ficheiro robots.txt para verificar se está a utilizá-lo corretamente com a ferramenta de análise do robots.txt disponível nas "Ferramentas do Google para *Webmasters*".
- Faça um esforço para garantir que os anúncios não interferem nas classificações dos motores de busca. Por exemplo, os anúncios do *Google AdSense* e os *links* de *DoubleClick* são impedidos de serem rastreados por um ficheiro robots.txt.
- Se a sua empresa adquirir um sistema de gestão de conteúdo, verifique se o sistema cria páginas e *links* que podem ser rastreados pelos motores de busca.

⁴⁸ Lynx é um *browser* de texto muito configurável para a *World Wide Web* (atualmente é o *browser* mais antigo em uso).

- Utilize o ficheiro robots.txt para evitar a indexação de páginas de resultados de busca ou outras páginas geradas automaticamente que não acrescentam valor para utilizadores provenientes de motores de busca.
- Teste o *site* para ter a certeza que ele aparece corretamente em diferentes *browsers*.
- Monitore o desempenho do seu *site* e optimize os tempos de carregamento. A meta do Google é providenciar aos utilizadores resultados mais relevantes e uma ótima experiência. *Sites* ágeis aumentam a satisfação do utilizador e melhoram a qualidade geral da *web* (especialmente para os utilizadores com conexões lentas de internet), e esperamos que, enquanto os *webmasters* aperfeiçoam os seus sites, a velocidade geral da *web* também aperfeiçoe.

3. Diretrizes de Qualidade

a. Princípios Básicos

- Crie páginas principalmente para os utilizadores, não para os motores de busca. Não engane os seus utilizadores nem apresente aos motores de busca um conteúdo diferente daquele que você exibe aos utilizadores, o que normalmente é referido como *Cloaking*⁴⁹ (camuflagem do conteúdo real da página).
- Evite truques que pretendem melhorar a classificação nos motores de busca. Uma boa regra prática é imaginar se você sentir-se-ia confortável explicando as suas ações para um *site* concorrente. Outro teste útil é perguntar a si mesmo: "Tal ia ajudar aos meus utilizadores? Eu faria isto se os motores de busca não existissem?"
- Não participe em esquemas de *links* que são usados para aumentar a classificação do seu *site* ou que manipulem o *PageRank*. Evite, principalmente, *links* para *spammers* ou para "más vizinhanças" na *web*, já que esses tipos de *links* podem prejudicar a classificação do seu *site*.
- Não use programas não-autorizados para enviar páginas, verificar classificações, etc. Esses programas consomem recursos de computação e

⁴⁹ Explica-se com mais detalhe na secção 2.4.2 como uma das técnicas *Black Hat* SEO.

violam os nossos Termos de Serviço. O Google não recomenda a aplicação de produtos como *WebPosition Gold* que envia consultas automáticas ou programadas ao Google.

b. Diretrizes específicas

- Evite usar texto ou *links* ocultos.
- Não utilize técnicas de *Cloaking* (camuflagem do conteúdo real da página) nem faça redirecionamentos não autorizados.
- Não envie consultas automáticas ao Google.
- Não carregue páginas com palavras-chave irrelevantes.
- Não crie várias páginas, subdomínios ou domínios com conteúdo duplicado.
- Não crie páginas com comportamento malicioso, como *phishing* ou instalação de vírus, troianos ou outros tipos de *softwares* nocivos.
- Evite usar páginas de entrada⁵⁰, criadas apenas em função dos motores de busca, ou outras técnicas "pré-fabricadas", como programas de filiação, que têm pouco ou nenhum conteúdo original.
- Se o seu *site* faz parte de um programa de filiação, certifique-se de que ele agrega valor. Forneça um conteúdo exclusivo e relevante que atraia a utilizadores a visitarem o seu *site* primeiro.

B. Yahoo

Yahoo divide as suas diretrizes conforme se segue [Yahoo 2012]:

1. Obtenção de uma classificação elevada para o seu *site*

- As listagens de busca orgânica no Yahoo são fornecidas pelo Bing. Familiarize-se com as ferramentas disponíveis na “Central do Bing para *Webmasters*”.
- Pense cuidadosamente nos termos-chave que o público irá usar para encontrar conteúdo como o seu. Use estes termos para orientar a construção das páginas e texto do seu *site*.

⁵⁰ Conhecida comumente como *doorway page*, é outra técnica *Black Hat* SEO. Ver mais detalhes na seção 2.4.2.

- Os utilizadores são mais propensos a clicar em um *link*, se o título corresponde à sua pesquisa. Escolha termos para o seu título de página que correspondam ao conteúdo dos seus documentos.
- Use o *Meta Tag* com o atributo *Description* e escreva a sua descrição com precisão e cuidado. Depois do título, a descrição é a característica mais importante para os utilizadores. Certifique-se de que o título e a descrição do documento chamam a atenção do utilizador e correspondem ao conteúdo do seu *site*.
- Use o *Meta Tag* com o atributo *Keyword* para listar palavras-chave em cada página do seu *site*. Use uma lista diferente de palavras-chave relacionadas a cada página específica no *site*. Não use um grande conjunto de palavras-chave para o *site* todo.
- Manter o texto e os *links* relevantes em HTML. Codificar o texto em gráficos ou mapas de imagem pode impedir que os motores de busca encontrem o texto ou sigam os *links* para outras páginas do seu site.
- Adicionar um *sitemap* HTML/XML, com um *link* de sua página de boas-vindas, pode ajudar a garantir que todas as suas páginas sejam rastreadas.
- Use o texto ALT para os gráficos. O texto ALT melhora o conteúdo de texto de sua página para fins de pesquisa. Isto também faz que as suas páginas sejam mais acessíveis para os *browsers* de somente texto e os visitantes de *site* com deficiência visual.
- Mantenha correspondência com *webmasters* de *sites* relacionados e outros provedores de conteúdo relevantes, a fim de aumentar o número de *links* para o seu site. NOTA: *Links* de *sites* semelhantes ao seu ou de sites cujo público poderia estar interessado no seu *site* irá melhorar a classificação de busca do seu site. No entanto, a construção de *Link Farms*, os quais criam *links* entre as páginas não relacionadas, a fim de aumentar na página a contagem de *link* não irá ajudar o seu *site*. O uso de *Link Farms* viola as diretrizes de *site* de Yahoo! e não irá a melhorar a classificação da sua página.

2. Diretrizes de qualidade de conteúdo

- Conteúdo original e único, de valor autêntico.

- Páginas destinadas essencialmente para pessoas, as considerações para os motores de busca ficam em um segundo plano.
- *HyperLinks*⁵¹ destinados para ajudar as pessoas encontrarem conteúdo relevante e interessante, quando for o caso.
- Metadados (incluindo título e descrição) que descrevem com precisão o conteúdo de um *site*.
- Bom *design* do *site* no geral.
- Páginas que o Yahoo! não deseja incluir no índice.
- Páginas que prejudicam a precisão, diversidade ou relevância dos resultados de busca.
- Páginas dedicadas ao redirecionamento do utilizador para outra página (páginas de entrada).
- Vários *sites* ou páginas que oferecem fortemente o mesmo conteúdo.
- *Sites* com inúmeros e desnecessários nomes virtuais de *hosts*.
- Páginas produzidas em grandes quantidades, que foram geradas automaticamente ou que são de pouco valor (páginas *cookie cutter*).
- Páginas que usam métodos para aumentar artificialmente a classificação no motor de busca.
- O uso de textos ou *links* que são ocultados do utilizador.
- Páginas que fornecem conteúdo diferente do que o utilizador final vê (*cloaking*).
- *Sites* com excessivas ligações cruzadas com outros *sites* para aumentarem a popularidade aparente de um *site* (esquema de *links*).
- Páginas desenvolvidas essencialmente para motores de busca ou páginas com palavras-chave excessivo ou fora de contexto.
- Uso impróprio de nomes de concorrentes.
- Vários *sites* que oferecem o mesmo conteúdo.
- *Sites* que usam muitas janelas *pop-up* que interferem na navegação do utilizador.
- Páginas que parecem duvidosas, fraudulentas ou oferecem uma experiência de utilizador de baixa qualidade.

⁵¹ Ou simplesmente *link*, é uma palavra, grupo de palavras, ou imagem que o utilizador clica para ir para um novo documento ou para uma nova seção dentro do documento atual.

C. Bing

O Bing tem focado as suas diretrizes nos itens a seguir [Bing 2012]:

1. Recomendações técnicas para o seu *site*

- Use apenas código HTML bem formatado nas suas páginas. Certifique-se que todos os *tags* emparelhados sejam fechados e que todos os *links* possam abrir a página correta. Para obter informações sobre validar o seu código HTML, consulte o Serviço de Validação de Marcação W3C ou use uma ferramenta semelhante.
- Se o seu *site* contém *links* quebrados, o Bing pode não ser capaz de indexar o seu site de forma eficaz, evitando assim que as pessoas atinjam a totalidade das suas páginas.
- Se mover uma página, crie o URL da página original para redirecionar as pessoas a nova página. Indique se o movimento é permanente ou temporário, e use o redirecionamento apropriado (use um redirecionamento 301 para uma mudança permanente ou um redirecionamento 302 para uma mudança temporária).
- Use um ficheiro robots.txt ou os *Meta Tags* para controlar quais páginas e que outros conteúdos não estão disponíveis no seu *site* para ser indexado por Bing e outros motores de busca. Você pode usar o ficheiro robots.txt para prevenir que os *crawlers* rastreiem ficheiros e diretorias específicas.
- Certifique se o seu ficheiro robots.txt está configurado para permitir que o *crawler* do Bing rastreie o seu *site* e não esteja na lista de *crawlers* proibidos de indexar o seu site.
- Mantenha os seus URLs simples e estáticos. Os URLs que são complicados ou que mudam com frequência são difíceis de indexar como *links* de destino. Por exemplo, o URL www.example.com/mypage é mais fácil para o Bing rastrear e para as pessoas digitar do que uma longa URL com múltiplas variáveis. Além disso, o URL que não muda ou usa múltiplas variáveis é menos complicado para as pessoas se lembrarem e marcarem. Isso torna a sua página um *link* de destino mais apropriado do que outros sites.

- Preste atenção para o *software* malicioso (*malware*). Se uma ou mais páginas do seu *site* contém ou liga diretamente para páginas externas que contêm *malware* ou conteúdo malicioso, como uma imagem maliciosa ou ficheiro de documento corrompido, ou um controle *ActiveX* ou *JavaScript* prejudicial, essas páginas serão desativadas e sublinhadas como *malware* nas SERPs do Bing.
- Registe o seu *site* no Bing Webmaster Tool. Um dos principais benefícios para o registo de um *site* é que a página *default* do site será sempre indexada.
- Outro benefício importante de registar o seu *site* no Bing *Webmaster Tools* é poder usar a ferramenta *Submit URLs* para pedir que o Bing indexe páginas específicas do seu *site* que não estão no índice. Enquanto as páginas que não usam técnicas *web spam*, nem contêm *malware* podem ser submetidas ao Bing através desta ferramenta e serem adicionadas ao índice sem nenhum problema. Note-se que esta ferramenta limita o número de submissões de URL para um máximo de 10 por dia e 50 por mês.
- Adicionar um ficheiro *sitemap.xml* para o seu *site*, o qual irá ajudar o Bing a encontrar todas as suas páginas relevantes. Os *links* que são incorporados em menus, caixas de listagem e elementos similares não são acessíveis aos *crawlers*, a menos que elas apareçam no seu *sitemap*.
- Mantenha a hierarquia do *site* bastante simples. Isto é, cada página só deve estar de uma a três cliques distantes da página *default*.

4. Diretrizes de conteúdo para o seu *site*

- No texto visível da página, inclua palavras-chave que os utilizadores possam escolher como termos de consulta de busca para encontrar a informação no seu *site*.
- Crie texto original, com ótimas palavras-chave para o elemento `<title>` e o *Meta Tag* com o atributo *Description* para cada página.
- Limite todas as páginas para um tamanho razoável. O Bing recomenda cobrir um tópico por página. O tamanho das imagens numa página HTML não deve ser menor do que 150 KB.

- Certifique-se de que cada página seja acessível por pelo menos um *link* de texto estático.
- Não coloque texto que quer indexar dentro de imagens. Por exemplo, se quiser que o nome da empresa ou a morada seja indexada, certifique-se que não seja exibida somente dentro de uma imagem do logótipo da sua empresa.

5. Técnicas que impedem que o seu *site* apareça nos resultados do Bing

- Tentar aumentar a densidade de palavras-chave de uma página pela adição de um grande volume de palavras irrelevantes. Isto inclui encher a etiqueta com o atributo ALT que os utilizadores provavelmente não vêem.
- Usar texto ou *links* ocultos. Utilize apenas texto e *links* que sejam visíveis para os utilizadores
- Usar técnicas, tais como *link farms*, para aumentar artificialmente o número de *links* para sua página.