

Università di Pisa
Laurea Magistrale in Informatica Umanistica

I MOTORI DI RICERCA NELL'ERA DELLA SEO

MARIATERESA MORETTI

Matricola: 526083

Anno Accademico 2014/2015

Sommario

Introduzione.....	3
Un po' di storia.....	4
SEO: definizione e significato.....	5
White Hat SEO vs Black Hat SEO.....	5
Black Hat SEO	6
White Hat SEO.....	6
Motori di Ricerca: come funzionano.....	7
Basi del design e dello sviluppo	9
Utilizzo e targeting delle parole-chiave.....	10
Aumentare la popolarità ed i link	11
La figura del “SEO Specialist”.....	11
False verità sui motori di ricerca	12
Errato utilizzo della SEO: come reagisce Google	14
Conclusioni.....	15
Bibliografia.....	16
Sitografia	16

Introduzione

Nel 1989 l'immagazzinamento, l'accesso e la ricerca di collezioni, di documenti e di informazioni di ogni genere, fu rivoluzionata da un'invenzione chiamata World Wide Web e dal suo fondatore Tim Berners-Lee.

L'invenzione di documenti collegati – linkati – fu veramente rivoluzionaria a quel tempo, nonostante Vannevar Bush, un tempo Direttore dell'ufficio “Resercher and development”, avesse prefigurato questa invenzione in un saggio del 1945, “As we may think¹”. Nel saggio descrive una macchina futuristica – molto simile al moderno pc – che rispecchia il processo cognitivo umano che lascia delle tracce di connessione tra le collezioni di documenti. Quattro decenni dopo, i resti della macchina ipotizzata da Bush, diventavano lo scheletro del web di Berners-Lee.

Il World Wide Web diventa il segnale ultimo della morte dell'era industriale e l'inizio del dominio dell'era dell'informazione. Eppure, nonostante la rivoluzione nell'immagazzinamento e accesso alle informazioni inaugurato dal web, gli utenti che iniziano ad usare il web per le ricerche si trovano a dibattere sullo stesso; essi si sentivano dispersi in una enorme libreria, come nella breve storia di Jorge Luis Borgers del 1941, “La Biblioteca di Babele”²

Molte delle informazioni nella “biblioteca del web” rimangono inaccessibili come quelle nella “biblioteca di Babele”. Effettivamente, all'inizio, molti dei motori di ricerca non alleviavano la frustrazione degli utenti: le ricerche dovevano essere condotte ordinando le informazioni gerarchiche degli argomenti di Yahoo! o setacciando le tante pagine web restituite dal motore di ricerca e cliccando sulle pagine per determinare, manualmente, quale fosse la più importante.

Oggi le cose sono cambiate notevolmente. I motori di ricerca sono diventati validi alleati nel reperimento delle informazioni. Quando si cercano informazioni di qualsiasi genere, per togliersi dubbi su qualsiasi argomento o se si vuole imparare qualcosa, cosa si fa oggi?

1. Si seguono corsi a pagamento, anche molto dispendiosi, che non sempre garantiscono reali conoscenze;
2. Si cercano informazioni su internet che, dalla sua diffusione ad oggi, sembra essere diventato il posto privilegiato della conoscenza e delle ricerche del mondo.

A ben guardare, i corsi con professionisti stanno velocemente lasciando il posto alla didattica on-line e *googlare*, il nuovo termine coniato il seguito all'affermarsi sempre maggiore di Google come maggiore motore di ricerca, ben racchiude quello che si è portati a fare oggi grazie anche alla disponibilità dei dispositivi mobili e alla possibilità di connettersi ad internet senza costi aggiuntivi.

Il meccanismo che scatta non appena una qualsiasi informazione viene *googlata* è quasi immediato: si cerca, indubbiamente e subito, il titolo più accattivante, meglio descritto e che ci illude e garantisce massima resa con minimo sforzo.

¹ Vannevar Bush . As we may think. Atlantic Monthly, 176(1):101-8,1945. (As we mat think, 1945)

² Jeroge Luis Borgers. *La biblioteca di Babele*. 1941

Il Team di Google, forte della sua posizione di leader mondiale dei motori di ricerca ha definito una serie di standard che permettono ad un sito di essere ai primi posti nel momento della ricerca.

Dalla loro ideazione ed implementazione nascono i SEO, acronimo di *Search Engine Optimization* ovvero letteralmente **Ottimizzazione per i Motori di Ricerca**, che hanno stravolto le regole di ricerca su internet ma anche di indicizzazione di un sito per le ricerche degli utenti.

Un po' di storia

Dopo l'invenzione del World Wide Web nel 1989, nasce nel 1990 ALIWEB³ (Archie Like Indexing on the Web), il primo motore di ricerca della storia. Verso la metà degli anni '90, momento di massima espansione di altri motori di ricerca come Lycos, AltaVista e Yahoo!, viene usato per la prima volta il termine *SEO* – Search Engine Optimization – anche se molti webmaster avevano già iniziato ad ottimizzare i siti senza saperlo.

Questo è un periodo, nell'evoluzione del web, in cui tutto è consentito per portare un sito nelle prime posizioni dei risultati di ricerca. I primi studi sugli algoritmi dei motori di ricerca si devono a tutti quei programmatori che, agli albori di internet, “smontavano e montavano” i risultati di ricerca per studiare e comprendere i meccanismi di posizionamento.

Inizialmente la SEO ha come unico obiettivo quello di creare pagine ottimizzate per i motori per scalare le classifiche. La decodifica degli algoritmi diventa molto sofisticata nel 1999. I motori di ricerca reagiscono ed iniziano ad impiegare **algoritmi multipli** per ciascun posizionamento. Anche le aziende cominciano ad adeguarsi a questi standard assumendo programmatori per *crackare* gli algoritmi dei motori di ricerca.

Iniziano le prime guerre aperte alle SEO, una vera caccia alle streghe. Se i siti contenevano anche solo una parola come “ottimizzazione” o “promozione” rischiavano di essere *bannati* dal motore di ricerca. Molte aziende SEO iniziano a scomparire e difficilmente qualcuna riuscì a sopravvivere e lo stesso successe alle liste dei clienti che avevano ottimizzato i siti.

Con l'avvento e l'affermarsi di Google cambiano le regole del gioco. Il colosso dei motori di ricerca introduce **spider di nuova generazione** che rilevano aspetti come il numero di keyword in un testo e il numero di link importanti in entrata che sono programmati per rilevare contenuti duplicati e nascosti.

Il sistema di **PageRank** di Google e i suoi nuovi algoritmi tagliano fuori la maggior parte delle tecniche per ingannare i motori di ricerca che hanno funzionato fino a quel momento.

Mentre gli algoritmi di Google diventano sempre più performanti, gli esperti SEO non rimangono a guardare; notano che i link in entrata hanno una grande importanza e trovano strategie innovative per procurarsi migliaia di link in modo automatizzato.

Emblematico è il caso scoppiato nel 2006 in Italia: vengono pubblicati molti articoli con link contenenti sempre la stessa parola. Ad esempio digitando “buffone” nella ricerca di Google appariva

³ <http://www.aliweb.com/>

al primo posto la pagina dell'allora Presidente del Consiglio. Per correggere i risultati Google ha dovuto modificare l'algoritmo⁴.

Ancora una volta il team anti-spam di Google si deve adattare alle nuove tattiche: Google e SEO ingaggiano una battaglia senza esclusione di colpi con vittorie e sconfitte ambo le parti.

Al giorno d'oggi Google domina incontrastato il regno dei motori di ricerca (seguito da *Bing* e *Yahoo!*) ed ha rivoluzionato l'approccio alla loro ottimizzazione.

La SEO ha subito a sua volta dei cambiamenti non indifferenti in seguito ai costanti aggiornamenti degli algoritmi che stravolgono le strategie di posizionamento e la costringono ad orientarsi in una nuova direzione in cui il contenuto diventa, finalmente, la cosa più importante sul web.

Contenuti di qualità e non duplicati, longevità dei domini, mappe del sito funzionali e tanti altri elementi, spingono la SEO verso aspetti qualitativi e non più ai sotterfugi e le astuzie del passato.

SEO: definizione e significato

Per SEO – *Search Engine Optimization* che significa letteralmente *Ottimizzazione per i motori di ricerca* – si intende quell'insieme di strategie pratiche volte ad aumentare la visibilità del proprio sito internet migliorandone la posizione nelle classifiche dei motori di ricerca all'interno nei risultati non a pagamento, detti risultati puri ed organici.

Queste pratiche sono molteplici e riguardano i diversi aspetti di un sito: l'ottimizzazione della struttura, del codice HTML, dei contenuti testuali, la gestione dei link in entrata ed in uscita. Poiché Google è il motore di ricerca più usato al mondo, la maggior parte delle attività SEO riguardano lo studio dell'algoritmo di Google e dei suoi aggiornamenti. La sfida della SEO è quella di riuscire a proporre all'utente per primi il risultato migliore, ovvero i risultati maggiormente pertinenti alla sua ricerca.

Fare SEO significa, prima di tutto, *scrivere contenuti ricchi e pertinenti*. Altra caratteristica importante della SEO è quella di cercare di ottenere un buon numero di link da siti il più possibile pertinenti e rilevanti; nel World Wide Web avere molti link al proprio sito significa aver costruito un sito importante sotto diversi punti di vista e che permette di rispondere in maniera esaustiva alle ricerche degli utenti.

White Hat SEO vs Black Hat SEO

Le tecniche utilizzate dai webmaster per garantire massima visibilità dei loro siti nei motori di ricerca sono: Black Hat SEO; White Hat SEO e Gray Hat SEO. Vediamole nel dettaglio.

⁴ <http://www.lastampa.it/2006/04/08/blogs/web-notes/google-bombing-contro-berlusconi-bKi9O2zQ3sFzBEEPSXbnyI/pagina.html>

Black Hat SEO

L'origine del nome risale al 1953, quando Robert Ebert, critico cinematografico, sottolineò come nei film western il cappello nero fosse sempre indossato – e quindi sinonimo – dal protagonista cattivo. Da qui la distinzione tra chi agisce secondo la legge e chi, invece, aggira il sistema.

La tecnica **Black Hat SEO** ha avuto un grande successo negli anni passati, quando gli algoritmi di ricerca erano ancora poco raffinati e venivano aggirati con facilità da web master per far risalire nel ranking delle SERP⁵ le pagine internet.

Quali sono le tecniche più usate dai *Black Hat Specialist*? La più comune è quella di ripetere più volte nel testo della pagine le **parole-chiave** per le quali si vorrebbe far trovare il testo; un'altra tecnica è quella di inserire nel testo anche dei **link nascosti** utilizzando un colore uguale a quello dello sfondo della pagina in modo da farli mimetizzare; oppure, si possono usare delle **doorway pages** ovvero pagine create con il solo scopo di spingere in alto nelle SERP altre pagine senza però fornire vero contenuto agli utenti. Ci sono anche Black Hat Specialist che si affidano al **link spam**, ovvero l'utilizzo di software per creare link a supporto di una pagina o, ancora, creare pagine da fornire agli spider quando approdano sul sito, delle pagine senza contenuti visibili ma che sono create ad uso e consumo degli spider⁶.

Sebbene l'utilizzo delle tecniche Black Hat SEO si rivelino molto efficaci, soprattutto nel raggiungimento immediato dello scopo – saltare ai primi posti al momento della ricerca –, è frequente che il risultato si riveli temporaneo e soggetto a variazioni improvvise dovuto a delle “tecniche di protezione” poste in opera dai motori di ricerca. In ultima istanza, i motori di ricerca utilizzano come modalità di protezione quella di abolire dalle liste di ricerca il sito incriminato per l'utilizzo di queste tecniche.

White Hat SEO

White Hat SEO è una pratica usata per migliorare le prestazioni di ricerca e per mantenerle in linea con i termini e le condizioni di un motore di ricerca.

La tecnica White Hat SEO, in contrapposizione con la tecnica Black Hat SEO, si riferisce ad ogni pratica che migliora le ricerche in una SERP mantenendo, comunque, l'integrità del sito e rimanendo entro i termini di servizio dei motori di ricerca. Queste tecniche includono termini come:

- Offrire contenuti e servizi di qualità;
- Usare parole-chiavi descrittive e molti *meta-tag*;
- Rendere il sito facilmente navigabile.

La tecnica del White Hat SEO è quella che viene, ad oggi, maggiormente apprezzata dal team di sviluppatori di Google in particolare perché è quella che gli permette di rimanere il leader nei motori di ricerca in quanto garante dei contenuti per i milioni di utenti che lo utilizzano ogni giorno.

⁵ *Search Engine Result Page* che tradotto letteralmente significa *pagina dei risultati del motore di ricerca*, è una qualunque pagina di Google che compare dopo che avete inserito il termine o i termini da cercare.

⁶ Gli **spiders** sono dei piccoli programmi che partendo da una singola pagina Web (solitamente dalla home page) riescono a navigare attraverso tutte le pagine ad essa collegate

Infine troviamo il **Grey Hat SEO** che utilizza metodi che si muovono in una zona grigia, evitando al sito delle penalizzazioni, ma non garantendo la miglior qualità dei contenuti, in quanto è interamente focalizzato a migliorare i posizionamenti.

Motori di Ricerca: come funzionano⁷

I motori di ricerca hanno 4 funzioni principali:

1. Scansione- *crawling*- del web;
2. Costruzione dell'indice
3. Calcolo della rilevanza e dell'ordine dei risultati delle ricerca
4. Presentazione dei risultati.

Il processo di *crawling e indicizzazione* serve ad esplorare ed indicizzare i miliardi di documenti, pagine, files, notizie, video e altri media presenti sul World Wide Web. I motori di ricerca hanno bisogno di un modo per esplorare tutte le informazioni presenti nel web e per permettere all'utente di raggiungerle con il miglior percorso possibile.

La struttura del web serve a creare collegamenti tra tutte le pagine esistenti. Attraverso i link, gli spider possono raggiungere miliardi di documenti interconnessi tra loro.

Quando i motori di ricerca trovano queste pagine analizzano l'indice che si trova in esse e ne archiviano le parti in grandi hard drive in modo da poterle richiamare quando risultano necessarie per una qualsiasi ricerca.

I *data center*, in cui sono archiviate le pagine in modo da essere accessibili in pochi secondi quando necessario, contengono, a loro volta, migliaia di server che processano una mole inimmaginabile di informazioni. D'altra parte quando un utente effettua una ricerca attraverso qualsiasi motore di ricerca si aspetta un risultato immediato e anche qualche secondo di ritardo può essere "fatale" e può lasciare l'utente insoddisfatto.

Quando l'utente fa la ricerca, i motori devono setacciare miliardi di pagine indicizzate e fare due cose:

- Restituire solo i risultati rilevanti o utili rispetto alla ricerca effettuata
- Ordinare quei risultati in base ad un criterio di qualità o di importanza.

Questo processo di ottimizzazione sui motori di ricerca ha lo scopo di influenzare sia la rilevanza che l'importanza dei risultati presentati.

La *rilevanza* significa molto più che presentare semplicemente una pagina che contiene in modo particolarmente evidente le parole utilizzate per la ricerca; inizialmente i risultati forniti non andavano oltre questo semplice controllo, con conseguenze sulla qualità dei risultati presentati. Oggi, attraverso una evoluzione iterativa, sono stato elaborati dei modi migliori per individuare i risultati di qualità apprezzabili dagli utenti.

Anche l'*importanza* dei risultati è un concetto difficile da quantificare ed i motori di ricerca lo fanno del modo migliore possibile. La maggior parte dei motori di ricerca identifica l'importanza di un sito

⁷ <http://www.neoseo.it/guida-seomoz/funzionamento-motori-di-ricerca.php>

con la sua *popolarità* associandola in maniera direttamente proporzionale ai contenuti: quanto più un sito è popolare tanto più alto è il valore dell'informazione presente nei contenuti.

Popolarità e rilevanza non sono determinati manualmente. I motori di ricerca utilizzano degli algoritmi matematici per separare i buoni risultati da quelli cattivi e quindi ordinarli. Questi algoritmi sono costituiti da migliaia di componenti che nell'ambito del Search Marketing si è soliti definire *ranking factor*. Sono noti pochi espedienti che permettano di raggiungere ottimi risultati e ottenere il miglior traffico. Le poche informazioni e le pratiche di ottimizzazione sono fornite dagli stessi motori di ricerca e non sono neanche esaustive. Alcune delle linee guida all'ottimizzazione dei principali motori di ricerca sono:

YAHOO!	BING	GOOGLE
<ul style="list-style-type: none"> - Numero di altri siti che linkano al sito - Contenuto delle pagine - Aggiornamenti apportati all'indice - Test di nuove versioni del prodotto - Scoperta di nuovi siti - Cambiamento dell'algoritmo di ricerca 	<ul style="list-style-type: none"> - Includere, nel testo visibile, le parole che gli utenti potrebbero scegliere come termini di ricerca - Limitare il peso delle pagine ad una dimensione ragionevole (usare un argomento per pagina) - Assicurarsi che ciascuna pagina sia accessibile da almeno un link statico - Non inserire il testo che si vuole indicizzare nelle immagini. 	<ul style="list-style-type: none"> - Progettare le pagine per gli utenti (non per i motori) - Creare un sito con una chiara gerarchia e link testuali. - Ogni pagina dovrebbe essere raggiungibile da almeno un link statico - Creare un sito utile e ricco di informazioni con pagine che descrivono in modo chiaro e accurato i contenuti del sito. - Assicurarsi che i tag <code><title></code> e gli attributi ALT siano descrittivi e precisi - Mantenere un numero ragionevole di link per ogni pagina

Un altro modo per scoprire come funzionano i motori di ricerca consiste nel fare un po' di ricerche e andare a guardare le specifiche dei requisiti inserite nei brevetti depositati dai principali motori di ricerca all'Ufficio Brevetti degli Stati Uniti. Il più famoso è il *PageRank* di Google depositato con il brevetto #6285999⁸. Attraverso lo studio dei brevetti, degli esperimenti e i test, i search marketer hanno imparato a conoscere molte delle operazioni base effettuate dai motori di ricerca ed i componenti critici per creare siti e pagine con migliori posizionamenti e maggior traffico.

Quindi riassumendo, le principali caratteristiche di un buon motore di ricerca sono:

⁸ <http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetacgi%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=6285999.PN.&OS=PN/6285999&RS=PN/6285999;>
<http://infolab.stanford.edu/~backrub/google.html>

1. Crawling e indicizzazione: i siti con una struttura di link imperfetta possono impedire ai motori di trovare tutti i contenuti del sito [ved. Fig.1], oppure, pur permettendo ai motori di trovarli lasciano i contenuti così poco esposti che vengono ritenuti poco importanti per i motori e quindi non si trovano ai primi posti nel ranking della SERP; le pagine web che fanno uso di Flash – e altri applicativi – hanno un contenuto che non tutti i motori di ricerca sono in grado di leggere.
2. Interpretazione dei contenuti testuali: il testo non presentato sotto forma di HTML all'interno della parte di codice analizzabile di una pagina risulta invisibile ai motori di ricerca e questo include anche il testo presentato tramite immagini, foto, video e plug-in.
3. Matching tra contenuti e ricerche: Il testo scritto senza le parole che gli utenti usano quando eseguono ricerche sui principali motori di ricerca non è utile ad una buona indicizzazione del sito; se una parola è più usata di un suo sinonimo è bene usare la parola più utilizzata nell'indicizzazione; la stessa attenzione è da porre per le lingue straniere e le sue sfumature internazionali – se si scrivono dei contenuti in una lingua straniera, oltre ad usare il corretto uso della parola, è bene anche usare eventuali sfumature dovute alle diverse pronunce ed usi.

Basi del design e dello sviluppo

Per essere elencato nei risultati dei motori di ricerca, il materiale disponibile sul sito deve essere in formato HTML testuale. Immagini, files vari, Applet Java ed altri contenuti non testuali sono virtualmente invisibili agli algoritmi di ricerca dei motori, nonostante i progressi della tecnologia dei motori di ricerca.

Il modo più semplice per assicurarsi che parole e frasi presentate nei siti siano visibili ai motori di ricerca consiste, quindi, nell'inserirle nel testo XML. Sono, comunque, disponibili metodi avanzati come:

1. Alle immagini in formato .gif, .jpeg, .png possono essere assegnati degli attributi *alt* che forniscono ai motori di ricerca una descrizione testuale del contenuto visivo;
2. Le immagini possono essere mostrate ai visitatori come sostituti del testo usando fogli di stile .CSS;
3. I contenuti inseriti in Flash o plug-in Java possono essere ripetuti in formato testuale nella pagina;
4. I contenuti multimediali (audio e video) dovrebbero essere accompagnati da una trascrizione testuale se si vuole che vengano indicizzati dai motori.

I motori di ricerca hanno bisogno di vedere i contenuti per poter inserire le pagine nei loro indici basati su parole-chiave. Hanno anche bisogno di accedere ad una struttura di link esplorabile, una struttura che permetta agli spider di muoversi attraverso i percorsi del sito e che consenta loro di trovare tutte le pagine di un sito. Moltissimi siti commettono l'errore di nascondere o rendere difficoltosa la navigazione del sito per i motori di ricerca, influenzando in questo modo l'opportunità di vedere le proprie pagine indicizzate dai motori di ricerca.



Fig.1 come si comporta uno spider nel momento in cui raggiunge un sito con link visibili

In Fig.1 viene mostrato come si comporta uno spider quando raggiunge un sito; se, come detto, tutti i contenuti non sono resi in formato testo, non tutte le pagine che compongono il sito verranno correttamente visualizzate e quindi il posizionamento nella ranking del sito in questione sarà basso perché questo non sarà ritenuto capace di soddisfare la richiesta dell'utente. Nell'esempio viene analizzato lo spider di Google che ha raggiunto la pagina "A" e vede i link alle pagine "B" ed "E". Tuttavia, nonostante le altre pagine "C" e "D" possono essere pagine importanti del sito, lo spider non ha modo di raggiungerle (o anche solo di sapere se esistono) perché nessun link diretto e raggiungibile punta a queste pagine. Agli occhi di Google queste pagine risultano inesistenti.

Utilizzo e targeting delle parole-chiave

Le parole-chiave sono fondamentali nel processo di ricerca ed infatti, tutti gli algoritmi di recupero delle informazioni si basano su queste poiché i motori leggono ed indicizzano i contenuti delle pagine web e ne tengono traccia in indici basati su parole-chiave. In questo modo, invece di archiviare 25 miliardi di pagine web in un database, i motori hanno milioni e milioni di database più piccoli, ognuno focalizzato su una particolare parola-chiave o frase. Ciò aiuta i motori a rendere molto più veloce il recupero dei dati di cui hanno bisogno in una frazione di secondo.

Le parole-chiave dominano anche le nostre intenzioni di ricerca e l'interazione con i motori. Un modello di ricerca comune potrebbe funzionare in questo modo:

Quando la ricerca viene effettuata, il motore sa quali pagine recuperare in base alle parole inserite nel box di ricerca. Altri dati come l'ordine delle parole, l'ortografia, i segni di punteggiatura, l'uso di lettere maiuscole forniscono ulteriori informazioni che i motori possono utilizzare per affinare le ricerche ed ordinarle propriamente. Per ovvie ragioni, i motori di ricerca misurano i modi in cui le parole-chiave sono utilizzate nelle pagine per determinare la rilevanza di un particolare documento rispetto alla ricerca effettuata. Uno dei modi migliori per ottimizzare il posizionamento di una pagina consiste nell'assicurarsi che le parole-chiave siano ben viste nei titoli, nel testo e nei meta dati.

Aumentare la popolarità ed i link

Per ogni motore di ricerca che esplora il web, i link costituiscono le strade che collegano ciascuna pagina. Attraverso l'analisi dei link, i motori possono scoprire la relazione tra le diverse pagine.

I complicati algoritmi di analisi sono capaci di valutare un sito o una pagina in base alle informazioni che riescono a cogliere dall'analisi dei link che puntano al sito valutato o che escono dal sito valutato. I professionisti della SEO ritengono che tra le variabili utilizzate dai motori dei propri algoritmi, quelle relative ai link abbiano un peso particolarmente significativo. I motori analizzano la popolarità di una pagina e di un sito sulla base del numero e della popolarità delle pagine che linkano ad essi, oltre a metriche quali *trust*, *spam* e autorevolezza della pagina. I siti con maggiore *trust*⁹ hanno tendenza di norma a linkare a siti con altrettanto *trust*, mentre i siti con *spam*¹⁰ ricevono pochi link da siti dotati di elevato *trust*.

Nonostante gli algoritmi dei motori di ricerca siano in grado di calcolare il valore di ciascun link con valori matematici estremamente precisi, è impossibile per tutti coloro al di fuori del loro team di ingegneri conoscere quali siano realmente questi valori. Per tale motivo, la SEO si affida ad un numero di segnali in grado di guidarli nella costruzione di una scala di valori dei diversi link.

Oltre ai segnali dei link poco misurabili visti precedentemente; la figura della *SEO Specialist* è proprio quella figura che, grazie a molta pratica ed esperienza, risulta essere padrone di queste metriche e riesce a relazionarle al traffico ottenuto dai motori di ricerca.

La figura del “SEO Specialist”

In questo quadro in costante evoluzione – i motori di ricerca cambiano spesso i loro algoritmi –, come si inserisce la figura dello specialista della SEO?

È una figura professionale che si occupa di migliorare la visibilità del sito nei motori di ricerca ma anche di garantire che tale visibilità non sia frutto di sotterfugi ma sia dovuta all'importanza del sito, ai contenuti che esso contiene e tratta ed al tipo di informazioni che questo può fornire all'utente.

Importante per uno specialista è che non sia un semplice programmatore web ma possa contare su caratteristiche e competenze che vanno oltre la mera programmazione; è importante che sia capace leggere le specifiche dei “manuali” definiti dai vari motori di ricerca. È necessario, quindi, che sia un analista di testi in senso critico per garantire contenuti efficaci per gli utenti.

Ovviamente la SEO non può essere considerata una scienza esatta anche e soprattutto perché non si può prevedere quello che verrà richiesto esattamente in fase di ricerca dall'utente (*Google Instant*¹¹ cerca di contenere il “libero arbitrio” dell'utente quasi condizionandolo, in fase di digitazione sulla

⁹ Il *trust* (traduzione letterale “fiducia”; la traduzione concettuale sarebbe “affido” intendendo l'affido di beni mobili/immobili) è un istituto del sistema giuridico anglosassone di *common law*, sorto nell'ambito della giurisdizione di *equity*, che serve a regolare una molteplicità di rapporti giuridici di natura patrimoniale

¹⁰ Lo *spamming*, detto anche fare *spam* o *spammare*, è l'invio di messaggi indesiderati (generalmente commerciali o offensivi). Può essere attuato attraverso qualunque sistema di comunicazione, ma il più usato è Internet, attraverso messaggi di posta elettronica, chat, tag board, forum, Facebook e altri servizi di rete sociale.

¹¹ <http://www.google.it/intl/it/insidesearch/features/instant/about.html>

barra di ricerca, nelle sue scelte suggerendo le parole da cercare. Ma non è questo argomento di questo testo) e perché i trend di ricerca cambiano continuamente e senza preavviso. Lo Specialist SEO deve essere sempre aggiornato sia per gli algoritmi di ricerca sia per le priorità dei motori di ricerca; non cedere alle facili lusinghe di metodi per aggirare i controlli nei motori perché non è mai una tecnica vincente che permette di ingannare l'onnisciente team di sviluppo e controllo di un motore di ricerca.

False verità sui motori di ricerca

Nel corso degli ultimi anni che hanno visto la costante diffusione del web, si sono diffuse false verità sul funzionamento dei motori di ricerca e su ciò che serve per poter avere successo nel *search marketing*.

Alla fine degli anni '90 e ai tempi della SEO "classica", cioè i tempi in cui le ottimizzazioni per le ricerca iniziavano a diffondersi e a diventare una disciplina importante in ambito informatico, i motori di ricerca avevano un modulo per segnalare il proprio sito che costituiva parte del processo di ottimizzazione dello stesso. I proprietari del sito ed i webmaster associavano al sito e alle pagine informazioni (anche parole-chiave) e le inviavano ai motori che le scansionava e le includeva in pagine del proprio indice.

Dal 2001, la segnalazione del proprio sito ai motori di ricerca non solo non è più necessaria ma è anche in teoria inutile. I motori hanno tutti pubblicamente sottolineato che usano solo raramente la loro lista di URL e che il metodo migliore è quello di ottenere link da altri siti in modo da esporre i propri contenuti in modo naturale.

Si possono ancora vedere le pagine di segnalazione di Yahoo!, Google e Bing ma non trovano spazio nella SEO moderna.

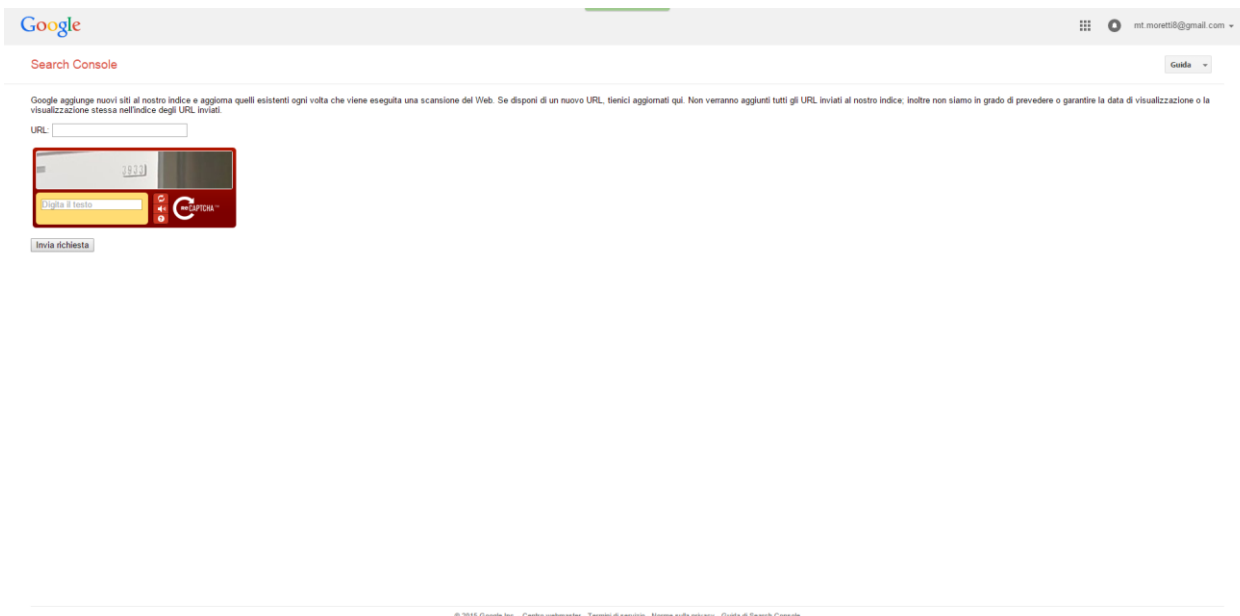


Fig.2 – modulo di segnalazione del sito¹² di Google–

¹² <https://www.google.com/webmasters/tools/submit-url>

Un'altra falsa idea riguardava i meta-tag che inizialmente rivestivano una grande importanza all'interno del processo di ottimizzazione del proprio sito. Bastava inserire le parole-chiave per le quali si voleva posizionare il sito e quando gli utenti facevano una ricerca utilizzando quelle parole la pagina appariva tra i risultati. Questo processo è diventato una pratica usata per attività di spam e ad oggi solo Yahoo! tra i principali motori di ricerca tiene ancora conto dei mega-tag per l'indicizzazione dei contenuti.

Un mito duro a morire nel mondo della SEO ruota attorno al concetto di *densità di parola-chiave* che si ottiene dividendo matematicamente il numero di volte in cui ricorre una parola-chiave per il numero totale di parole di una pagina; questo valore dovrebbe essere utilizzato dal motore per determinare la rilevanza ed il posizionamento di una pagina e dovrebbe assorbire tutti gli sforzi SEO. Nonostante questo mito sia stato sfatato più e più volte, continua a sopravvivere come dimostrano i numerosi strumenti SEO costruiti attorno alla presunta importanza delle densità delle parole-chiave. Le parole-chiave vanno, però, usate con intelligenza ed in base all'usabilità del sito. Il falso valore generato dall'aggiungere 10 volte in più una parola (sembra essere tornati alle pratiche Blak Hat già descritte) è di gran lunga inferiore ad un solo link editoriale di qualità ottenuto da un sito "serio".

L'importanza di pubblicità a pagamento e l'opinione che questa possa aiutare a migliorare i risultati ed il posizionamento di un sito è la prima vera grande "teoria di cospirazione della SEO". Ma ciò non è mai vero. I tre principali motori di ricerca utilizzati – Google, Yahoo! e Bing – separano nettamente all'interno delle pagine di presentazione dei risultati, quelli che sono a pagamento (i cosiddetti SEA – *Search Engine Advertising*) da quelli che non lo sono. Fino a quando la separazione tra le cose rimarrà netta, si è certi credere che la pubblicità su motori di ricerca non darà alcun beneficio a quanti la utilizzeranno per aumentare la loro posizione all'interno dei risultati.

The image shows a Google search results page for the query "stagione metropolitan new york". The search bar at the top shows the query and the number of results (Cerca 234.000.000 risultati (0,35 secondi)). The results are divided into two main sections: SEA (Search Engine Advertising) and SEO (Search Engine Optimization).

SEA (Search Engine Advertising): This section is highlighted with a red box and labeled "SEA". It includes an advertisement for "Metropolitan Opera in NY" with a red arrow pointing to the text "Google AdWords". The ad text includes "Metropolitan Opera in NY", "Tickets for Metropolitan Opera at Lincoln Center in New York City", and "Visualizza i tuoi annunci qui". To the right of the ad, there is a list of bullet points:

- Attività di Advertising a pagamento
- La **visualizzazione** dei messaggi è **gratuita**
- **Si paga solo** per l'azione del consumatore – l'accesso al sito (click)
- Premiata Rilevanza dell'annuncio e del sito
- Posizioni assegnate tramite un'asta automatica

SEO (Search Engine Optimization): This section is highlighted with a green box and labeled "SEO". It includes organic search results for "Metropolitan Opera in NY", "Metropolitan Opera in NY", "Metropolitan Opera in NY", "Metropolitan Opera - Wikipedia", "Metropolitan Museum of Art - Wikipedia", "Metropolitano di New York - Wikipedia", "Welcome to the Metropolitan Opera Family of websites - The Metropolitan Opera", "Metropolitan Opera", "The Metropolitan Museum of Art - Home", "Metropolitan di New York, LA STAGIONE", "In cinema italiani stagione Metropolitan - Ansa - Libero 24x7", and "Metropolitan Opera di New York al via la nuova stagione | Ritz...". A green arrow points from the "SEO" label to the text "Risultati Organici:", which is followed by a list of bullet points:

- Generati da un **algoritmo**
- Imparziali e **non acquistabili**
- Premiata Rilevanza e Importanza

Fig.3 Differenza di visualizzazione tra SEO e SEA¹³ in una pagina di ricerca di Google

¹³ Immagine tratta dal sito: <http://www.studiosamo.it/wp-content/uploads/2013/01/differenza-seo-sea.jpg>

Errato utilizzo della SEO: come reagisce Google

Dalla sua diffusione ad oggi, la politica SEO non ha garantito solo, a quanti la utilizzano in maniera corretta, ottimi risultati in fatto di visibilità e “presenza” ai primi posti nel ranking della SERP, ma è stato anche causa di alcuni “problemi” che hanno portato i motori di ricerca a prendere provvedimenti nei confronti di quanti violano i requisiti fondamentali.

Oltre al caso già descritto del 2006 in Italia la letteratura on-line riporta altri casi in cui il colosso americano è dovuto correre ai ripari e, all’occorrenza, sanzionare i siti.

Nel 2006, il team di Google decise di *bannare* dalle liste il sito della BMW Germania ¹⁴. Le indagini condotte da Google hanno scoperto che il sito della casa automobilistica manipolava i risultati del motore di ricerca per far apparire tra i primi posti il proprio sito. I webmaster della BMW avevano implementato un doppio “attacco” scorretto all’algoritmo di Google; da un lato con la tecnica del Black Hat fornivano risultati diversi agli utenti rispetto alle loro ricerche, mentre dall’altro forniva ai motori di ricerca risultati ancora diversi rispetto a quelli forniti agli utenti con l’utilizzo di *doorway pages* per migliorare il posizionamento¹⁵.

Esempio lampante di utilizzo mancato rispetto delle linee guida definite da Google, il caso della BMW tedesca è da annoverare tra i casi in cui il colosso americano delle ricerche sul web non ha battuto ciglio anche al cospetto di una delle aziende leader nel settore automobilistico. Le politiche di controllo dei motori di ricerca non sono state scritte per favorire i siti di agenzie più importanti ma sono state definite per garantire all’utente la migliore esperienza di utilizzo del web. Dopo la sentenza di Google, il sito della BMW.de è stato riscritto ex-novo ed è stato poi riaccolto negli indici di ricerca.

Un altro esempio è il caso scoppiato riguardo la pagina di Inteflora.uk¹⁶. Nel 2013 la versione britannica di Google ha penalizzato il sito ufficiale dei Interflora. La penalizzazione ha visto il sito ufficiale dell’azienda di spedizione di fiori non apparire nelle prime posizioni della SERP pur digitando delle parole-chiave che potessero in maniera coerente rimandare all’azienda. In realtà, il sito appariva in prime posizioni ma nella parte della pagina relativa alla SEA e quindi non al sito ufficiale ma allo spazio pubblicitario comprato su Google. La motivazione principale che ha spinto Google potrebbe essere stata la politica di *bad linking* adottata dall’azienda dovuta ai collegamenti provenienti da blog di utenti o da altri portali.

Anche se in questo caso è stato preso in considerazione solo Google, esistono in letteratura altri casi in cui i motori di ricerca hanno eliminato dalle liste dei siti colpevoli di aver male utilizzato a proprio vantaggio i metodi di ottimizzazione.

¹⁴ <https://coteire.wordpress.com/2012/08/30/why-bmw-site-was-banned-by-google/>

¹⁵ Il posizionamento (nella sua accezione di attività) è l’insieme delle azioni volte a migliorare la posizione di una pagina web nelle pagine di risultato dei motori di ricerca (SERP).

¹⁶ <http://www.webnews.it/2013/02/22/google-e-interflora-non-ce-rosa-senza-spine/>

Conclusioni

In un mondo frenetico e sempre in crescita come quello di oggi anche l'arte di imparare e conoscere ha cambiato i suoi ritmi. Oramai non ci si ferma più tanto tempo per apprendere quello che ci richiede la nostra esperienza quotidiana e sembra che "tutto e subito" sia diventato il primo ed unico motto che anima le nostre giornate. Anche nella scelta di cosa imparare non ci si ritrova mai a vedere quello che, realmente, interessa la nostra "fama di sapere" ma ci ritrova a leggere quello che, più facilmente, ci garantisce di soddisfare la fame di sapere con il minimo sforzo e la massima resa.

In questo contesto di "conoscenza lampo", Internet svolge un ruolo importante e fondamentale; una grande biblioteca in cui trovare tutto quello di cui abbiamo bisogno, quando si vuole e dove si vuole.

I vari team dei motori di ricerca in tutto il mondo (Google è il primo ma non è l'unico) hanno garantito una forte indicizzazione del sapere e, con la creazione di specifici algoritmi di ricerca, hanno permesso a tutti gli utenti, sia agli esperti utilizzatori di Internet sia a quanti non sono proprio abituati ad usare questo tipo di tecnologia, di accedere alle informazioni richieste in poco tempo e con un'attinenza ai contenuti cercati molto alta.

Cosa ci riserva l'ottimizzazione per i motori di ricerca?

Dare visibilità ai contenuti che gli utenti stanno cercando e che rispondano esattamente ai loro bisogni. Questa è la relazione più proficua che ci si possa aspettare tra i motori di ricerca e la SEO.

Di conseguenza l'ottimizzazione legittima, spesso non sarà neanche visibile pur rimanendo trasparente, perché si concentrerà a migliorare le prestazioni, l'architettura, l'usabilità e i contenuti veicolando i valori di marketing e diventando parte integrante del prodotto e dell'esperienza dell'utente.

La SEO dovrà affrontare un approccio multidisciplinare in cui tutti gli aspetti della tecnologia, Search Engine Marketing, Social Media Marketing e Content Marketing dovranno integrarsi in versioni ottimali di Web Marketing.

Bibliografia

As we mat think. (1945). In V. Bush. Atlantic Monthly.

la Biblioteca di Babele. (1941). In J. L. Borgers.

Google's pagerank and beyond the science of search engine ranking. Amy N. Langville and Carl D. Meyer

Sitografia

(n.d.). Retrieved from Aliweb:

<http://www.aliweb.com/>

(n.d.). Retrieved from The Anatomy of a Large-Scale Hypertextual Web Search Engine:

<http://infolab.stanford.edu/~backrub/google.html>

differenza seo sa. (n.d.). Retrieved from studiosamo: Method for node ranking in a linked database:

<http://www.studiosamo.it/wp-content/uploads/2013/01/differenza-seo-sea.jpg>

Funzionamento motori di ricerca. (n.d.). Retrieved from NetSEO:

<http://www.netseo.it/guida-seomoz/funzionamento-motori-di-ricerca.php>

Google bombing contro Berlusconi. (n.d.). Retrieved from La Stampa:

<http://www.lastampa.it/2006/04/08/blogs/web-notes/google-bombing-contro-berlusconi-bKi9O2zQ3sFzBEEPSXbnyI/pagina.html>

google e interflora non c'è rosa senza spine. (n.d.). Retrieved from Webnews:

<http://www.webnews.it/2013/02/22/google-e-interflora-non-ce-rosa-senza-spine/>

Method for node ranking in a linked database . (n.d.). Retrieved from Uspto Patent Full-Text and Image Database:

<http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnethtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=6285999.PN.&OS=PN/6285999&RS=PN/6285999>

Search Console. (n.d.). Retrieved from Google:

<https://www.google.com/webmasters/tools/submit-url?pli=1>

informazioni su Google Instant. (n.d) Retrieved from Google/dentro la ricerca:

<http://www.google.it/intl/it/insidesearch/features/instant/about.html>