
Google - Search Engine Optimisation

Sapere come e cosa Google cerca all'interno del suo indice, può aiutare un ricercatore esperto a formulare meglio le query.

Le informazioni contenute in questa sezione sono soggette a cambiamenti repentini senza preavviso. Prima di mettere in pratica questi consigli è bene cercare di verificare quali siano le attuali voci di corridoio su come Google crei il suo indice. In ogni caso, il confine tra il Search Engine Optimization (SEO) e cercare di barare, è molto fine. Visto che la pena è l'esclusione dagli indici di Google, il mio primo consiglio è di cercare di stare sempre dalla parte della ragione.

Una considerazione da fare quando si parla di SEO è che questo, molte volte, contrasta con i nostri sforzi di web-usability. Applicare quindi questi consigli di SEO con il dovuto buon senso.

Le informazioni contenute in questa sezione sono tutte frutto dell'esperienza personale. L'utilizzo di queste tecniche ha portato un quadruplicarsi del traffico di pc-facile.com dopo un solo mese dall'applicazione. Non sono in nessun ordine particolare in quanto si presuppone che chi affronta il problema del SEO abbia già tentato alcune delle tecniche qui elencate.

Alcune delle soluzioni proposte in questa sezione richiedono una conoscenza abbastanza approfondita di argomenti come HTML e web-server; si consiglia di fare dei back-up prima di apportare qualunque modifica.

Il tag <title>

Le informazioni contenute nel tag <title></title> sono molto importanti. Cercate di indicare chiaramente il contenuto della pagina con un titolo appropriato che contenga al suo interno le keyword importanti. Va da sé che i titoli delle pagine dovrebbero essere diversi per ogni pagina anche quando usate contenuti dinamici.

Se offrite una sezione per il download del software, avere un titolo uguale per tutte le schede software non vi aiuterebbe molto. Potrebbe essere più utile invece indicare il titolo del software insieme a qualche keyword. Su pc-facile.com abbiamo optato per dei titoli molto descrittivi come indicato qui sotto:

DivX [Download software Codec Microsoft]

Dove "DivX" è il nome di un codec (un particolare tipo di software), "Download software" dice all'utente cosa si può fare e allo stesso tempo sono due keyword molto ricercate su Internet, "Codec" è la categoria sotto cui è classificato DivX (ed è un'altra keyword importante) e Microsoft è la piattaforma su cui gira DivX (ed è una keyword).

Seguendo la medesima logica è importante che i titoli del contenuto che scrivete siano descrittivi, magari anche a discapito del sensazionalismo che potrebbero provocare. Un titolo come "Importante notizia" potrebbe attrarre molti utenti (ok, in realtà no, ma rende l'idea), ma a Google servirebbe a poco. Meglio qualcosa di più specifico e più ricco di keyword utili a Google.

L'HTML dei titoli

Il titolo di una pagina, di una sezione o di un paragrafo (inteso propriamente come titolo non come il tag <title>) viene solitamente indicato usando una combinazione di tag , , <i> e <u>. Questo va benissimo - come avremo modo di vedere più avanti -, ma non è la soluzione ottimale.

L'HTML prevede sei tag specifici per indicare un titolo (heading) e questi sono: <h1>, <h2>, <h3>, <h4>, <h5>, <h6> dove <h1> è il titolo più importante e l'importanza diminuisce nei seguenti tag.

Non posso sottolineare abbastanza l'importanza di usare questi tag, in particolar modo il tag <h1>. Fatelo.

Se i tag <h1...6> non soddisfano il vostro gusto estetico potete modificarne tutti gli aspetti facilmente usando i CSS.

Inserite il seguente codice nell'header (adattando font, dimensione... alle vostre esigenze). Includo qui anche come modificare un link all'interno dei tag <h1...6>, in quanto è stato un problema che mi ha portato via molto tempo.

```
<style type="text/css">
<!--
h1
{
font-family:Verdana, Arial, Helvetica, sans-serif;
font-size: 12pt;
color: #000000;
text-decoration: none;
margin-bottom: 0;
}
h1 a
```

```
{  
  font-family: Verdana, Arial, Helvetica, sans-serif;  
  font-size: 12pt;  
  color: #000000;  
  text-decoration: none;  
  margin-bottom: 0;  
}  
-->  
</style>
```

Fatto questo usate il tag `<h1...6>` esattamente come fareste per con gli altri tag.

```
<h1></h1>
```

Per ulteriori informazioni su come usare i CSS vi consiglio:

<http://www.w3schools.com>

[Nota: l'istruzione "margin-bottom: 0;" è tecnicamente inutile ma risolve un brutto problema estetico presente solo in alcune versioni di Internet Explorer.]

Dominio, nome e posizione

Il tipo e il nome del dominio sono fattori molto importanti. Google dà molta importanza ai domini di secondo livello e tende a considerare poco i domini di terzo livello. Se volete una presenza importante su Internet è vitale che il vostro dominio non sia del tipo `http://egdisegno.xoom.com`, ma sia invece `http://www.egdisegno.eu`

Altro fattore importante nella scelta del dominio è il nome. Mentre le regole di usability impongono che il nome sia breve, facilmente ricordabile e possibilmente fonetico, Google pare apprezzare i nomi di domini che contengono delle keyword come ad esempio:

<http://keyword1-keyword2-keyword3.com>

A voi il delicato equilibrio tra usability e la sempre crescente importanza dei motori di ricerca.

Nome della pagina

Il nome del dominio non è l'unico di cui Google tiene conto. I nomi delle pagine individuali sono molto importanti. Ad esempio la pagina `software.php?id=123`

dice a Google solo che si tratta di software. Invece la pseudo-pagina `software_DivX_codec.php` da molte informazioni sul contenuto della pagina. Questo tipo di lavoro è facile quando un sito è piccolo, ma quando si lavora con siti dinamici che rappresentano le pagine usando dei template fissi la cosa diventa difficile da gestire.

In nostro aiuto vengono il `mod_rewrite` su Apache e alcuni filtri su IIS di Microsoft.

Per fare in modo che al posto di questa pseudo-pagina, che non esiste, venga mostrata quella giusta senza tradire il nostro segreto è necessario inserire il seguente codice nel file `.haccess` di Apache (attenzione che diverse versioni di Apache possono differire leggermente: leggere quindi prima la documentazione).

```
RewriteEngine on  
RewriteBase /  
RewriteRule ^software_DivX_codec/?$ /software.php?id=25 [L]
```

Se usate invece IIS avete molte soluzioni per ottenere questo stesso risultato. Forse la più semplice, ma la meno potente consiste nell'installare la libreria URL Replacer messa a disposizione gratuitamente dallo sviluppatore cecoslovacco Antonin Foller.

<http://www.motobit.com/help/url-replacer-rewriter/iis-mod-rewrite.asp>

In testa è meglio

Google considera importante il contenuto che si trova in cima ad una pagina, quindi una buona tecnica è posizionare il contenuto importante il più in alto possibile. Il problema e la soluzione sembrano banali: non lo sono.

Prendiamo il caso tipico in cui la pagina è divisa in due colonne; dove la colonna di sinistra rappresenta la barra di navigazione, mentre la colonna di destra rappresenta il contenuto.

Il codice HTML che costruirebbe questa pagina (solo la parte che ci interessa) sarebbe:

```
<table>  
  <tr>  
    <td>codice per barra di navigazione</td>  
    <td>contenuto della pagina</td>  
  </tr>  
</table>
```

Come vedete il contenuto si trova dopo la navigazione. Ma con un piccolo espediente possiamo far salire il contenuto della nostra pagina al di sopra della navigazione senza influenzare la usability.

```
<table>
  <tr>
    <td>&nbsp;</td>
    <td rowspan="2">contenuto della pagina</td>
  <tr>
    <td>codice per barra di navigazione</td>
  </tr>
</table>
```

Nel caso, invece, abbiate la navigazione in alto il codice da usare è il seguente. Questo snippet non funziona perfettamente e, in alcuni casi, può dare dei problemi.

```
<table>
  <tr>
    <td>&nbsp;</td>
    <td rowspan="3" valign="bottom">contenuto della pagina</td>
  </tr>
  <tr>
    <td colspan="2">codice per barra di navigazione</td>
  </tr>
  <tr>
    <td>&nbsp;</td>
  </tr>
</table>
```

Contenuto, keyword, font, meta-tag e sessioni

Il segreto sta nel contenuto

Sembra una cosa banale da dire, ma un buon sito si basa su di un buon contenuto che viene aggiornato di frequente. La logica che sta dietro a questo ragionamento è semplice: più contenuto avete, più pagine create e ognuna di queste pagine è un potenziale accesso per gli utenti. Se il contenuto è fatto bene (sia dal punto di vista della qualità, sia dal punto di vista del SEO) ogni pagina non solo sarà un potenziale accesso, ma sarà anche ottimizzato rispetto ad alcune keyword.

Ma c'è di più, il contenuto che produceate - articoli, recensioni, tutorial... - dovrebbe essere distribuito su quanti più siti possibili. In cambio potete chiedere che in fondo al contenuto ci sia un link al vostro sito. Questa semplice tecnica, se eseguita bene, vi permetterà di ottenere un buon numero di link che contribuiranno

Keyword density

Uno dei metodi usati da Google per misurare l'importanza di una keyword all'interno di una determinata pagina è quello di calcolare la densità delle keyword. In altre parole: più volte ripetete una keyword e più Google penserà che quella keyword è rilevante all'interno di quella pagina. Anche qui Google sembra dare maggior peso alle keyword che si trovano in cima al testo.

Un avvertimento: cercate di non sovra-ottimizzare una pagina, il vostro primo obiettivo è che questa venga apprezzata dall'utente. Google viene sempre secondo.

Font

Google riconosce la dimensione del testo e dà importanza alle keyword scritte più in grande oppure in grassetto o sottolineate. Considerando che le keyword sono utili, sia all'utente durante la lettura che a Google, si consiglia di evidenziare le keyword importanti con il tag .

Inoltre, dove non potete usare il tag <h1...6> ma vi serve identificare un titolo o delle keyword, fate in modo che il titolo sia grande, anche a discapito della grafica della pagina. È stato notato che pagine contenente del testo esageratamente grande - e intendo tanto grande da riempire una pagina solo con il titolo - ottengono una rilevanza per quelle keyword ben al di sopra della media. Ovviamente usate il vostro buon senso.

Meta tag

Le voci sui meta tag sono piuttosto discordanti. Molti pareri però indicano che Google considera "description" e "keyword" abbastanza poco, se addirittura li considera. Se avete molto tempo, e volete coprire tutte le opzioni, lavorate sui meta tag, ma fate in modo di NON utilizzare un meta tag unico per ogni pagina del sito. Stando alle leggende il tag "description" è quello su cui lavorerei se proprio ne avessi il tempo. Altri motori, pare, diano molta importanza a questi tag.

```
<meta name="description" content="" >
```

```
<meta name="keywords" content="" >
```

Sessioni

Alcuni siti utilizzano un sistema di sessioni principalmente per cercare di garantire l'autenticazione dei propri utenti - ad esempio sui forum. Questo avviene aggiungendo in fondo all'URL una stringa alfanumerica valida per la durata di una visita e che però cambia ad ogni visita dell'utente.

I link

Al massimo 100 link

Non chiedetemi il perché, ma Google, nelle sue FAQ, indica esplicitamente che preferisce pagine che contengano meno di 100 link. Dato che non vogliamo contrariare Google: facciamolo e basta.

Per contare il numero di link in una pagina io uso questo semplice JavaScript. Per azionarlo è sufficiente inserirlo nella barra degli indirizzi e premere invio, oppure potete salvarlo nei vostri bookmark e richiamarlo quando serve.

Tanti link, per favore

Come vedrete nella parte dedicata a come Google calcola il PageRank, l'idea originale di Brin e Page era che una pagina era tanto più importante quanti più link puntavano a quella pagina (avevano applicato a Internet il concetto delle citazioni sulle pubblicazioni scientifiche). La matematica alla base dell'applicazione di questo concetto non è semplice, ma come vedrete, non sono importanti solo il numero di link, ma anche il PageRank che questi link hanno.

Nonostante alcune voci di corridoio suggeriscano che il PageRank non abbia più una grande influenza sul posizionamento di un sito all'interno dei risultati,

indizi più convincenti indicano il contrario. Ma che sia il PageRank a influenzare il posizionamento oppure sia una variante a farlo, a noi non interessa; ciò che ci interessa è che sembra certo che i link che puntano ad un sito conferiscono un punteggio che viene utilizzato nel calcolo del posizionamento.

Costruire, quindi, un solido network di link che puntano al proprio sito è molto molto importante anche se si rivelerà altrettanto faticoso e dispendioso in termini di tempo.

I punti da tenere a mente quando si vuole costruire questo network sono:

Creare una struttura di link che cresce gradualmente nel tempo. Pare Google riconosca quando tanti (troppi) link appaiono tutti in una volta sola e penalizzi questa pratica. Se potete aggiungete un nuovo link ogni giorno. Più i siti sono grandi e importanti e più alto sarà il loro PageRank. Cercate di farvi linkare da questi siti. Meglio ancora se questi siti sono rilevanti al vostro contenuto. Farvi linkare da siti piccoli non vi darà invece alcun beneficio. Se è possibile, e ritenete il costo ragionevole può essere una buona idea quella di pagare per essere inclusi in directory molto considerate. Cercate di fare in modo che il vostro link si trovi su di una pagina che non ha molti altri link in uscita; questo indebolirebbe il PageRank che vi viene trasmesso.

Verificate che il formato dell'URL sia sempre lo stesso. Gli esempi qui di seguito portano un utente alla stessa pagina, ma per Google si tratta di sei siti diversi:

<http://vostro-sito.com>
<http://vostro-sito.com/>
<http://vostro-sito.com/index.html>
<http://www.vostro-sito.com>
<http://www.vostro-sito.com/>
<http://www.vostro-sito.com/index.html>

Cercate di convincere chi vi linka a includere delle keyword per voi importanti all'interno dell'URL. Fate però in modo che tutti i link siano uguali, cercate di avere un po' di varietà: toglierà quella sensazione che i siti siano stati preparati meccanicamente.

I link reciproci sembra stiano perdendo favore all'interno di Google. Il vostro scopo è quello di ottenere link senza che voi dobbiate linkare a vostra volta. Specialmente non perdetevi troppo tempo a linkare pagine preposte solo allo scambio di link.

Evitate di linkare pagine che non sono negli indici di Google o di Yahoo. Evitate anche di linkare link farm, web-ring o qualunque altro tipo di sito che miri ad approfittarsi delle debolezze negli algoritmi di Google. Verreste black-listati e perdereste qualunque beneficio acquisito.

Anchor text

I link da siti esterni, verso il nostro sito, sono importanti, ma questi link vanno sfruttati meglio. Un link come questo dice poco a Google:

Scaricate DivX da qui.

Farsi invece linkare in questo modo è molto prezioso:

Cliccate qui per

scaricare DivX.

La logica dietro questo ragionamento è che i siti esterni saranno sempre più onesti riguardo al reale contenuto di una pagina di quanto possiate esserlo voi.

Ora immaginate di abbinare quanto detto prima agli anchor text. Otterremmo il seguente risultato e vi posso assicurare che sarebbe molto importante questo messaggio per Google.

Cliccate qui per

scaricare DivX.

Il consiglio è quindi non solo di cercare di avere tanti link che puntano al nostro sito, ma di cercare di influenzare il testo data la grande importanza che questo ricopre.

Link verso pagine esterne

Come vedrete nella parte dedicata a come Google calcola il PageRank, ogni link che porta dal vostro sito ad un sito esterno scala una piccola parte del vostro PageRank. Per limitare questo danno necessario cercate di limitare i link verso l'esterno e inserirli in pagine ricche di link verso altre pagine del vostro stesso sito.

Non fraintendetemi, però, i link verso pagine esterne sono importantissimi: un sito privo di link esterni viene cancellato dall'indice di Google! Ma anche non

considerando questo fatto, i link sono la base di Internet: un sito privo di questi link non è possibile che fornisca un servizio adeguato ai propri utenti.

Pagine lontane dalla root

Come regola di base considerate che Google dà maggiore importanza a quelle pagine che sono nella root directory - questa è la cartella in cui c'è la home page di un sito. A ogni pagina che si trova in una sotto-cartella della root viene tolto un punto di PR. Questo punto può essere recuperato con un buon cross linking o se è linkata da siti esterni, ma parte comunque con uno svantaggio.

Dato che non c'è un numero massimo di pagine che possono risiedere nella root, se non per un problema di organizzazione, si consiglia di tenere lì quante più pagine possibili.

Cross linking

Come vedrete nella parte dedicata a come Google calcola il PageRank, è possibile influenzare il modo in cui è distribuito il PageRank all'interno di un sito modificando la struttura dei link interni. La cosa non è facile, ma se avete necessità di dare importanza ad alcune vostre pagine potete provare diverse strutture finché non trovate quella che più fa al vostro caso. Per i vostri esperimenti potete usare gli script che trovate nell'appendice.

I primi link

I primi link, esterni al vostro sito, che potete ottenere facilmente sono anche quelli piuttosto ignorati dai novelli webmaster. Google dà una certa importanza a siti che sono indicizzate in Yahoo e in Open Directory. Seguite i link qui sotto e assicuratevi di essere presenti:

<http://www.dmoz.org/add.html>

<http://it.docs.yahoo.com/info/aggiungere.html>

Aggiungerei a questi anche altri due motori piuttosto snobbati, Altavista e Arianna. Non faranno granché ma vi concederanno un link in più:

<http://it.altavista.com/addurl>

<http://arianna.libero.it/addurl/inserisci.html>

E, ovviamente, non dimenticatevi di dire a Google di venire a cercarvi!

<http://www.google.com/addurl.html>

Consigli finali

Cose da NON fare!

Come già spiegato la pena per chi bara è l'esclusione dall'indice. Considerato quanto lavoro c'è dietro a un sito, il rischio dell'eliminazione dall'indice di Google è troppo elevato. Ecco qui la lista nera delle cose da NON fare:

NON usate del testo nascosto per aumentare la densità delle keyword senza disturbare l'utente

Tecniche comuni includono usare testo dello stesso colore dello sfondo, layer invisibili, frame nascosti...

Questo punto è particolarmente importante. Qui di seguito vi mostrerò come vengono realmente messi in pratica questi trucchi: ve lo mostro solo per farvi capire, NON usate questi mezzi!

I tag HTML usati per inserire commenti all'interno di una pagina sono un buon esempio. Sono nati per permettere l'introduzione di brevi annotazioni. Non essendo visibili agli utenti che visitano la pagina web vengono spesso abusati per far trovare ai motori di ricerca un po' di keyword in più rispetto a quelle mostrate agli utenti:

```
<!-- Questo è un commento in cui potrebbero esserci molte keyword -->
```

Nei siti che fanno uso di frame, il tag `<noframes>` è stato ideato per mostrare testi a quei browser che non possiedono il supporto dei frame o che hanno temporaneamente disabilitato tale caratteristica. Un chiaro abuso del tag `noframes` consiste nel riempirlo di keyword:

```
<noframes>  
Keyword visibili solo ai browser che non supportano i frame.  
</noframes>
```

Il tag `<noscript>` è a volte usato per fornire contenuti alternativi a quei browser che non supportano il linguaggio indicato nel corrispondente tag `<script>`. Quello che è nato per meglio supportare gli utenti dei browser testuali o dei browser meno aggiornati viene spesso utilizzato come un ulteriore trucco per fornire ai motori di ricerca testi non visibili alla maggior parte degli utenti:

```
<noscript>  
Testo visibile solo ai browser che non supportano il linguaggio indicato in un precedente tag SCRIPT.
```

</noscript>

Infine vorrei indicarvi un abuso del tag <alt> che viene utilizzato per descrivere le immagini o per mostrare del testo alternativo quando queste non sono visibili.

```
<img src="" alt?"Testo visibile durante un mouseover solitamente riempito di keyword">
```

Non riesco a ripetere a sufficienza di NON usare questi sistemi: imparateli, ma non utilizzateli!

Si pensa che Google stia utilizzando strumenti di Latent Semantic Indexing (LSI or LSA - Latent Semantic Analysis) che leggono un testo e permettono di determinare se questo è coerente o meno.

NO a pagine gateway o cloaking

Questa tecnica consiste nel creare delle pagine appositamente ottimizzate per i motori di ricerca che vengono date in pasto solo ai bot e non agli utenti veri.

NO link farm

Tecnica che prevede la creazione di un numero elevato di domini tutti interlinkati tra di loro in modo che questi possano ottenere un PageRank elevato. Questo PR può servire esclusivamente internamente, oppure, come spesso fanno società di SEO, serve per aumentare il PR dei clienti.

NO shadow domains

Tecnica simile al cloaking, ma al posto di una pagina si ottimizza un intero dominio. Il traffico viene poi convogliato con redirect invisibili al sito finale che deve ricevere gli utenti.

Google mi ha cancellato dall'index

Nel caso dovesse succedere il peggio contattate Google per chiedere di essere ri-inclusi o per vedere se ci sono dei problemi. È inutile che sottolinei l'importanza di essere educati e gentili in questo frangente.

help@google.com

Non sovra-ottimizzare

Il vostro primo obiettivo è che l'utente apprezzi il vostro lavoro, non che lo apprezzi Google. I canoni del posizionamento cambiano di continuo, cercare di stare al passo con Google è impossibile. E anche se ci riusciste, il giorno in cui cambiano questi parametri, vi ritrovereste in guai seri. Lavorate bene, ma tenete bene in mente che il vostro obiettivo finale è l'utente.

A dimostrazione di quanto sto dicendo si pensi al 16 Novembre 2003 quando Google, senza preavviso, modificò il modo di ordinare i risultati. All'aggiornamento, come agli uragani Statunitensi, fu dato un nome: Florida. La conseguenza fu che molti siti sparirono dalle prime posizioni mentre altri scalarono le vette. Molti siti che avevano ottimizzato le pagine scomparvero, così come molti siti di e-commerce, e, a oggi, non si sa ancora cosa sia successo esattamente.

Alcuni hanno speculato che sia stata una mossa per togliere visibilità ai siti di e-commerce e costringendoli a aderire al programma pubblicitario di AdWords. Altre teorie sostenevano che Google filtrasse le pagine troppo ottimizzate. La teoria più accreditata sembra, però, essere quella che Google abbia implementato un sistema detto "Hilltop"; cioè una tecnica in cui alcuni siti vengono indicati come esperti del settore e link da questi siti acquistano un'enorme importanza nel calcolo dell'ordinamento dei risultati.

Il 16 Novembre 2003 fu un duro colpo per molti siti di e-commerce. Chi aveva puntato molto sulla conoscenza dei meccanismi di Google si ritrovò improvvisamente senza la terra sotto i piedi. Non ripetete lo stesso errore.

Contrariamente a quanto sostengono molti, e a quanto possa sembrare da questo capitolo, l'unica fonte di documentazione affidabile è questa: <http://www.google.it/webmasters/>

Estratto da:

<http://www.pc-facile.com/>