



**UNIVERSITÀ DI PISA**

**Corso di Laurea in Informatica Umanistica**

**Ottimizzare un sito web per i motori di ricerca  
attraverso l'implementazione delle tecniche SEO**

**Candidato:** *Claudia Cinquini*

**Relatore:** *Maria Simi*

**Correlatore:** *Mirko Tavosanis*

Anno Accademico 2010-2011

<b>1. Introduzione .....</b>	<b>5</b>
<b>2. Premessa sulla SEO .....</b>	<b>6</b>
2.1 Come funziona un motore di ricerca.....	7
2.1.1 Analisi.....	8
2.1.2 Indicizzazione.....	8
2.1.3 Risposta .....	9
2.2 Criteri di indicizzazione per i motori di ricerca.....	12
2.3 Importanza del posizionamento nei risultati delle ricerche.....	13
2.4 Vantaggi della SEO.....	14
2.5 Sfide della SEO.....	16
<b>3. Il processo di ottimizzazione SEO .....</b>	<b>18</b>
3.1 Ricerca.....	18
3.1.1 Settore di riferimento.....	19
3.1.2 Analisi competitiva.....	19
3.1.3 Stato del sito .....	19
3.1.4 Parole chiave .....	20
3.2 Pianificazione strategica.....	20
3.2.1 Contenuti.....	20
3.2.2 Sviluppo dei link.....	20
3.2.3 Piattaforma tecnologica .....	20
3.3 Implementazione .....	21
3.3.1 Ottimizzazione interna.....	21
3.3.2 Ottimizzazione esterna .....	21
3.4 Monitoraggio .....	21
3.5 Valutazione e manutenzione .....	22

<b>4. Strategie per le parole chiave: analisi e ricerca.....</b>	<b>23</b>
4.1 La scelta delle parole chiave corrette .....	23
4.1.1 La teoria delle <i>Long Tail</i> .....	23
4.1.2 Parole chiave e lessico.....	24
4.2 Gli strumenti utili per le parole chiave.....	26
<b>5. Ottimizzazione interna on-page.....</b>	<b>29</b>
5.1 Head tags.....	29
5.1.1 Tag Title .....	30
5.1.2 Meta tag Description .....	31
5.1.3 Meta tag Keyword .....	31
5.1.4 Meta tag Robots.....	32
5.2 Body tags.....	32
5.2.1 Heading tag.....	33
5.2.2 Anchor text .....	33
5.2.3 Le immagini e il tag Alt.....	35
5.3 Densità delle parole chiave .....	37
5.4 Prominenza delle parole chiave.....	38
5.5 Prossimità delle parole chiave.....	38
<b>6. Ottimizzazione interna on-site.....</b>	<b>40</b>
6.1 Le URL.....	40
6.2 Struttura del sito .....	42
6.2.1 Usabilità e accessibilità del sito.....	44
6.2.2 Architettura dei link.....	46
6.3 Sitemap XML e robots.txt.....	47
<b>7. Ottimizzazione off-site .....</b>	<b>51</b>
7.1 Popolarità .....	51
7.1.1 Link esterni.....	52

7.1.2 PageRank .....	53
7.1.3 TrustRank .....	54
7.2 Come aumentare la popolarità.....	56
7.2.1 Social bookmarking.....	57
7.2.2 Feed RSS .....	58
7.2.3 I contenuti .....	59
7.2.4 I social network .....	60
<b>8. Caso di studio: www.toscanaper.me.....</b>	<b>61</b>
9.1 CMS Wordpress .....	62
9.2 Analisi delle parole chiave .....	64
9.3 Ottimizzazione .....	66
9.4 Analisi delle statistiche.....	68
<b>9. Conclusioni .....</b>	<b>71</b>
<b>Appendice A .....</b>	<b>73</b>
<b>Riferimenti bibliografici .....</b>	<b>74</b>

# 1. Introduzione

I motori di ricerca rappresentano a oggi uno dei principali strumenti di ricerca e conoscenza, visto che il 91% di italiani che naviga sul web, ritiene che siano lo strumento più efficace per cercare informazioni, prodotti e servizi<sup>1</sup>.

La ragione del loro utilizzo così ampio, probabilmente deriva dalla semplicità e dalla rapidità con cui i motori di ricerca riescono a farci trovare ciò che cerchiamo. Qualsiasi informazione si voglia cercare, infatti, basta aprire il proprio browser, digitare alcune parole sull'argomento che ci interessa trovare, e in pochi secondi, "si aprono" di fronte a noi una serie di pagine che possono essere in grado di soddisfare le nostre richieste.

In una realtà così vasta e in continua espansione come il web però, a volte accade che i risultati della nostra ricerca non corrispondano alle nostre aspettative. Le cause possono essere diverse: partendo dalle ragioni più ovvie, bisogna dire che non tutti gli argomenti reperibili sul web hanno lo stesso volume di interesse e di ricerca da parte degli utenti. Quindi, ad esempio, un argomento specialistico o fortemente di nicchia, non sarà trattato da molti siti, e di conseguenza, è possibile non trovare tra i risultati della ricerca ciò che si stava cercando. Altre cause di questo disagio, riguardano invece aspetti più tecnici. Può dipendere da una scarsa, o assente, attenzione da parte dei webmaster nell'osservazione delle regole per consentire ai motori di ricerca una corretta indicizzazione, durante la fase di progettazione e realizzazione di un sito e delle sue pagine. Oppure, può dipendere dall'utilizzo di espedienti poco corretti, che aiutano il posizionamento di un sito nei motori di ricerca ma che sacrificano la qualità dei risultati. In ogni caso i motori di ricerca, e in particolare Google, lavorano costantemente per migliorare i propri algoritmi di indicizzazione, per renderli più sensibili e soddisfare l'utente con risultati nelle ricerche sempre più accurati e precisi. E' importante quindi per i webmaster, saper costruire siti web che offrano contenuti di qualità e che tali contenuti siano visibili dai motori di ricerca.

Questo è il difficile compito che si prefigge la *SEO*, acronimo che sta per *Search Engine Optimisation* (ottimizzazione per i motori di ricerca), e che è definita come "l'insieme di

---

<sup>1</sup>Sems.it, *Gli italiani e i motori di ricerca*,2010 [<http://www.sems.it/>]

tutte le attività finalizzate ad aumentare il volume di traffico che un sito web riceve dai motori di ricerca”<sup>2</sup>.

Scopo del mio elaborato, è quello di illustrare le principali tecniche e strategie in grado di ottimizzare un sito web per i motori di ricerca, dimostrando, attraverso un caso di studio, come sia possibile aumentare il traffico proveniente dai motori di ricerca su un blog, appena nato, realizzato dall’azienda presso cui ho svolto il tirocinio. L’obiettivo finale, è l’aumento del numero di visite al sito e il raggiungimento di un posizionamento nelle prime pagine dei risultati di ricerca di Google. Per raggiungere questo scopo, non sono state utilizzate campagne per la promozione del sito con annunci Pay Per Click, ma sono state implementate solamente le tecniche SEO. Il ruolo che ho ricoperto in questo progetto, è stato nella ricerca e nello studio delle keyword per l’ottimizzazione on-page, attraverso l’utilizzo dei principali strumenti per la ricerca delle keyword messi a disposizione gratuitamente da Google.

---

<sup>2</sup> Wikipedia, voce *Ottimizzazione (motori di ricerca)*  
[[http://it.wikipedia.org/wiki/Ottimizzazione\\_%28motori\\_di\\_ricerca%29](http://it.wikipedia.org/wiki/Ottimizzazione_%28motori_di_ricerca%29)]

## 2. Premessa sulla SEO

Prima di affrontare nel dettaglio le strategie di ottimizzazione che aiutano il posizionamento di un sito nei motori di ricerca, è giusto chiarire alcuni importanti concetti che stanno alla base della SEO.

### 2.1 Come funziona un motore di ricerca

Tutti i motori di ricerca, archiviano nei propri database le pagine web per identificare le potenziali risposte alle *query* poste dagli utenti, e lo fanno basandosi sulle tecniche dell'Information Retrieval<sup>3</sup>, una disciplina che si occupa della rappresentazione, della ricerca e della manipolazione di grandi collezioni di testi elettronici o di altri dati sul linguaggio umano<sup>4</sup>. L'architettura di base del sistema dell'Information Retrieval viene illustrata nella figura 2.1 che segue:

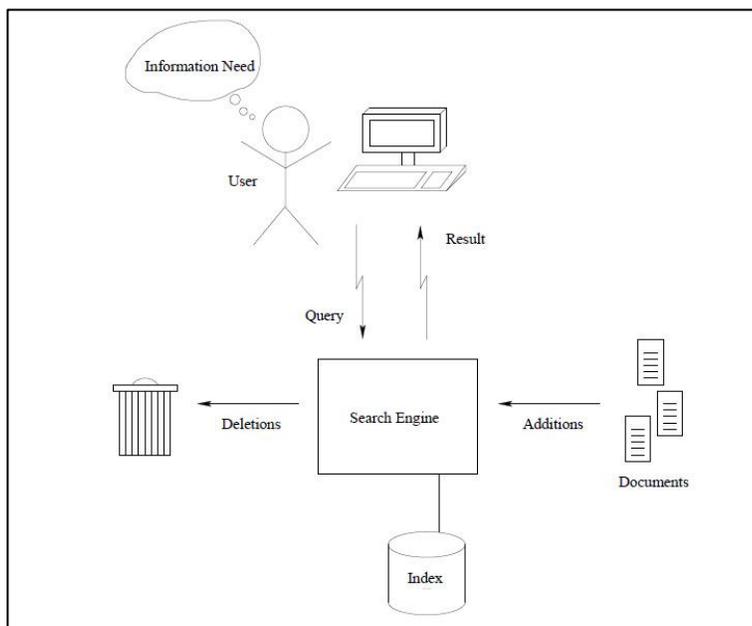


Figura 2.1 Architettura di base del sistema dell'Information Retrieval

<sup>3</sup> Abhishek Das, Ankit Jain, "Indexing the World Wide Web: The Journey So Far", *Next Generation Search Engines: Advanced Models for Information Retrieval*, 2011 [http://static.googleusercontent.com/external\_content/untrusted\_dlcp/research.google.com/it/pubs/archive/37043.pdf]

<sup>4</sup> Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, 2010 [http://www.ir.uwaterloo.ca/book/01-introduction.pdf]

Il processo di recupero delle informazioni su cui si basa un motore di ricerca si può dividere in tre fasi generali: analisi, indicizzazione e risposta.

### 2.1.1 Analisi

Questa prima fase di analisi si basa su software programmati per setacciare la rete, in modo da individuare costantemente nuovi siti e nuove pagine da indicizzare: si tratta degli *spider*, detti anche *robot*, *bot* o *crawler*. Gli spider selezionano in modo automatizzato i siti da visitare, cercando di raccogliere il maggior numero possibile di documenti per costruire i propri indici<sup>5</sup>. Per produrre risultati precisi e aggiornati, i motori di ricerca memorizzano una copia della pagina analizzata, detta *copia cache*. Lo spider, con una cadenza periodica, aggiorna la copia cache e la confronta con la pagina web, in modo che, se rileva modifiche significative, la pagina vecchia viene sostituita con quella nuova<sup>6</sup>. La frequenza con cui uno spider visita una pagina, dipende dalle sue dimensioni e dalla frequenza di aggiornamento dei contenuti.

### 2.1.2 Indicizzazione

Se la pagina viene considerata idonea dagli spider, viene inserita nel database del motore di ricerca, che, a sua volta, crea un indice della pagina. Questi indici, rappresentano la struttura principale dei dati che vengono utilizzati dai motori di ricerca per la ricerca e la classificazione della pertinenza dei risultati<sup>7</sup>. Se l'indice mancasse, infatti, per recuperare delle informazioni sulla base di una query, occorrerebbe una scansione completa di tutta la collezione, richiedendo molto tempo. In questa fase, quindi, i motori di ricerca provvedono alla creazione, alla conservazione e all'aggiornamento costante di un indice per l'intera collezione dei documenti presenti nel database. In questo modo, data una query, il motore si limita a scansionare solo l'indice, ottimizzando notevolmente i tempi

---

<sup>5</sup> Wikipedia, voce *Crawler* [<http://it.wikipedia.org/wiki/Crawler>]

<sup>6</sup> Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, 2010 [<http://www.ir.uwaterloo.ca/book/01-introduction.pdf>]

<sup>7</sup> Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, 2010 [<http://www.ir.uwaterloo.ca/book/01-introduction.pdf>]

di recupero dei documenti<sup>8</sup>. Per far questo, si utilizza un indice inverso della collezione, dove per ogni elemento  $w$  del vocabolario  $V$  considerato, viene indicato l'insieme delle occorrenze di  $w$  nella collezione di documenti e all'interno del documento stesso. I termini  $w$  sono elementi lessicali, ovvero singole parole o espressioni composte che vengono considerate come blocchi unici, mentre il vocabolario  $V$  è composto dai termini del documento rimasti dopo la fase preliminare di *processing*, che consiste nella *normalizzazione*, nello *stemming* e nell'eliminazione delle *stop words*<sup>9</sup>. Attraverso la normalizzazione, il documento viene trasformato in un elenco di token, e, con una serie di operazioni, vengono analizzati aspetti come l'utilizzo degli apostrofi e le contrazioni dei termini, l'eliminazione della punteggiatura e degli spazi. Lo stemming o lemmatizzazione, invece, è un processo algoritmico utilizzato nell'elaborazione del linguaggio digitale, che consiste nel determinare in modo automatico il lemma di una determinata parola, facilitando così le espansioni delle interrogazioni dei motori di ricerca<sup>10</sup>. Le stop words, infine, sono quelle parole che ricorrono molto frequentemente in un documento (congiunzioni, articoli e altre parole molto comuni) che però non forniscono informazioni utili per capire gli argomenti trattati in un determinato documento, perciò vengono eliminate, permettendo di risparmiare una notevole quantità di memoria.

### 2.1.3 Risposta

Utilizzando l'indice inverso, la raccolta di statistiche e un'altra serie di dati, il motore di ricerca è in grado di accettare la query digitata dall'utente nell'apposita casella di testo, e fornire un elenco dei risultati ordinato per grado di rilevanza<sup>11</sup>. La rilevanza è calcolata prendendo un insieme di pagine web che contengono i termini della query, che vengono analizzate sulla base di diversi parametri e classificate mediante un punteggio.

---

<sup>8</sup> Antonino Freno, *Indicizzazione di documenti di testo*  
[[http://www.dii.unisi.it/~freno/files/didattica/slides\\_indexing.pdf](http://www.dii.unisi.it/~freno/files/didattica/slides_indexing.pdf)]

<sup>9</sup> Abhishek Das, Ankit Jain, "Indexing the World Wide Web: The Journey So Far", *Next Generation Search Engines: Advanced Models for Information Retrieval*, 2011  
[[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/en//pubs/archive/37043.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en//pubs/archive/37043.pdf)]

<sup>10</sup> Wikipedia, voce *Stemming* [<http://it.wikipedia.org/wiki/Stemming>]

<sup>11</sup> Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, 2010 [<http://www.ir.uwaterloo.ca/book/01-introduction.pdf>]

Successivamente, i motori eliminano le pagine duplicate e ridondanti, e, infine, generano una sintesi delle pagine rimanenti, restituendo all'utente un elenco di link<sup>12</sup>. Questo elenco ordinato di risultati che restituisce il motore di ricerca è meglio noto con il termine *SERP* (*Search Engine Results Page*), che tradotto letteralmente significa “pagina dei risultati del motore di ricerca”<sup>13</sup>. La Figura 2.1 mostra un esempio di SERP su Google dopo aver effettuato una ricerca:

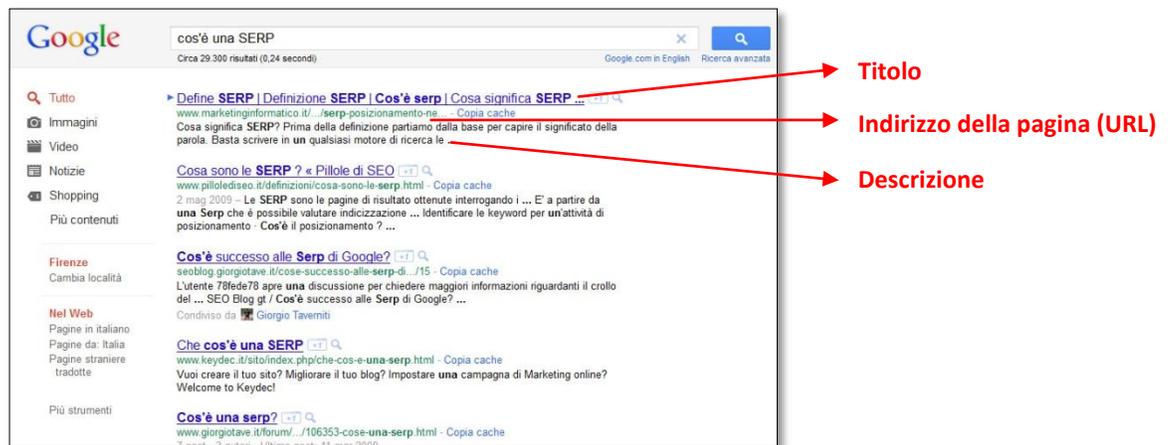


Figura 2.1 Esempio di SERP su Google

I criteri con cui viene calcolata la pertinenza di una determinata pagina sono di primario interesse per chi si occupa di SEO, ma i motori di ricerca non li svelano mai del tutto, perché su di essi si fonda l'intero business del web. Inoltre, i motori di ricerca non producono tutti gli stessi risultati, perché ogni motore di ricerca si basa su propri algoritmi di indicizzazione e posizionamento. Per queste ragioni, è fondamentale conoscere i fattori che incidono sul posizionamento di un sito nei risultati delle ricerche, anche se è impossibile ottimizzare un sito per tutti. Bisogna infatti decidere su quale motore di ricerca puntare, sulla base delle esigenze del sito e della preferenza degli utenti nell'utilizzo dei motori di ricerca.

Su Internet i motori di ricerca più utilizzati sono Google, Yahoo! e Bing, almeno nel mondo occidentale. In linea di massima comunque, è una buona scelta progettare un sito intorno a Google, come dimostrano i seguenti dati statistici: secondo uno studio condotto

<sup>12</sup> Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack *Information Retrieval: Implementing and Evaluating Search Engines*, 2010 [http://www.ir.uwaterloo.ca/book/01-introduction.pdf]

<sup>13</sup> Wikipedia, voce *Search Engine Results Page* [http://it.wikipedia.org/wiki/Search\_engine\_results\_page]

nel 2010 da SEMS e FullResearch<sup>14</sup> su un campione di 30.000 utenti Internet e 2000 intervistati, è emerso che in Italia Google ha una quota di mercato pari al 92%, mentre secondo i dati di comScore<sup>15</sup> di Maggio 2011, la quota di mercato Google negli Stati Uniti è del 65,4%. Yahoo! e Bing vengono utilizzati più negli USA, dove hanno rispettivamente il 16,9% e il 14,1, contro il 3% di Yahoo! e l'1% di Bing in Italia. Nelle tabelle 2.1 e 2.2 sono riportati i grafici relativi a questo studio, dove si può vedere la distribuzione completa dell'utilizzo dei diversi motori di ricerca in circolazione.

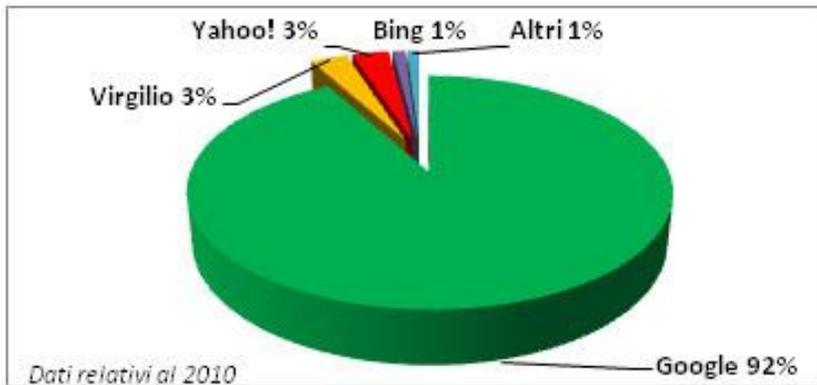


Tabella 2.1 Utilizzo dei motori di ricerca in Italia

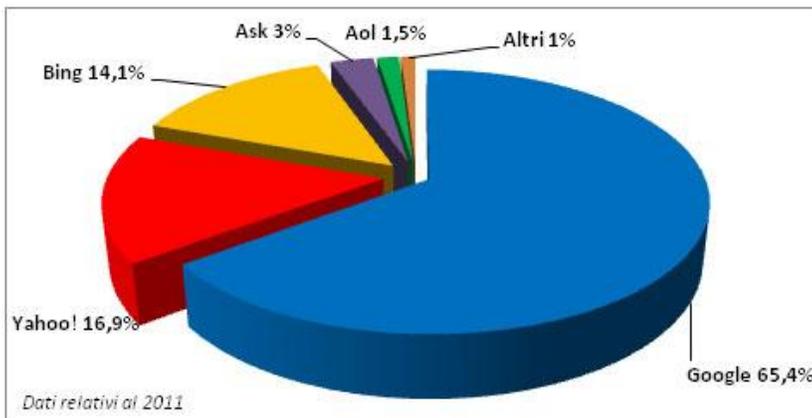


Tabella 2.2 Utilizzo dei motori di ricerca negli Stati Uniti

<sup>14</sup>Riccardoperini.com, *Statistiche di utilizzo dei motori di ricerca*, 2010  
[\[http://www.riccardoperini.com/statistiche-utilizzo-motori-di-ricerca.php\]](http://www.riccardoperini.com/statistiche-utilizzo-motori-di-ricerca.php)

<sup>15</sup>ComScore.com, *comScore Media Metrix Ranks Top 50 U.S. Web Properties for May 2011*  
[\[www.comscore.com/Press\\_Events/Press\\_Releases/2011/6/comScore\\_Media\\_Metrix\\_Ranks\\_Top\\_50\\_U.S.\\_Web\\_Properties\\_for\\_May\\_2011\]](http://www.comscore.com/Press_Events/Press_Releases/2011/6/comScore_Media_Metrix_Ranks_Top_50_U.S._Web_Properties_for_May_2011)

## 2.2 Criteri di indicizzazione per i motori di ricerca

Come si è detto, non è possibile ottimizzare un sito per tutti i motori di ricerca. In ogni caso, è possibile individuare i principali criteri che i motori di ricerca prendono in considerazione per il posizionamento, che sono illustrati nelle tabelle 2.3 e 2.4, e relativi allo studio realizzato da SeoMoz<sup>16</sup> nel 2009 e nel 2011.

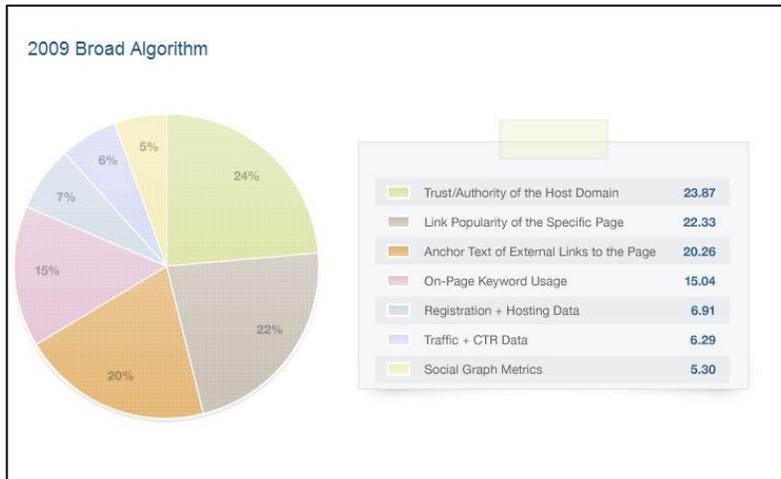


Tabella 2.3 Fattori che gli algoritmi prendono in considerazione per il posizionamento nel 2009

Nel 2009 i principali aspetti in grado di influire sul posizionamento riguardavano: l'autorità e la fiducia relative al dominio (23,87%), la link popularity di una determinata pagina (22,33%), gli anchor text nei link esterni (20,26%) e l'uso delle keyword all'interno della pagina (15,4%).

Confrontando i dati con quelli attuali, si nota un cambiamento nella valutazione dei criteri di posizionamento. I quattro aspetti più influenti adesso riguardano: le metriche relative al numero di link che rimandano a una pagina (21,45%), l'autorità di un determinato dominio (21,13), l'utilizzo delle parole chiave nelle pagine (14,93%) e nel nome del dominio (10,73%).

<sup>16</sup> SeoMoz.org, *2011 Search Engine Ranking Factors* [<http://www.seomoz.org/article/search-ranking-factors#predictions>]

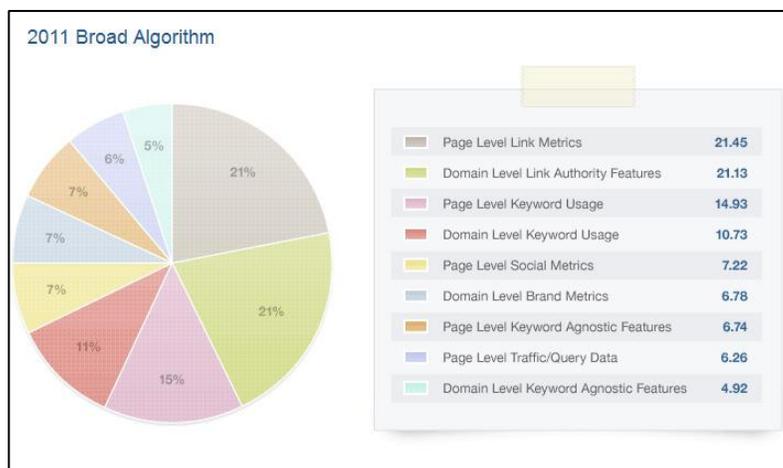


Tabella 2.4 Fattori che gli algoritmi prendono in considerazione per il posizionamento nel 2011

Alla luce di queste informazioni, si può dire che i motori di ricerca stanno affinando i loro algoritmi per fornire agli utenti siti sempre più autorevoli e risultati sempre più vicini alle loro richieste. Per chi si occupa di SEO, invece, questi dati servono a capire che, per aiutare il posizionamento di un sito, si deve puntare sulla giusta scelta delle keyword con le quali si vuole far indicizzare il sito dai motori, e lavorare per far crescere la fiducia e l'autorità del sito e delle sue pagine agli occhi dei motori di ricerca, proponendo contenuti che possano essere utili e originali, in modo da guadagnare in modo naturale link su altri siti.

## 2.3 Importanza del posizionamento nei risultati delle ricerche

Un sito che non sia ben indicizzato e visibile ai motori di ricerca è un sito inutile, perché difficilmente verrà trovato, e di conseguenza è come se non esistesse. Il posizionamento di un sito nelle pagine dei risultati di ricerca rappresenta quindi un aspetto cruciale. Infatti, come illustrato nella tabella 2.5, secondo uno studio realizzato da SEMS e FullResearch<sup>17</sup> su un campione di 30.000 utenti Internet italiani e 2000 intervistati, è emerso che il 32% degli utenti, quando effettuano una ricerca con un motore, si

<sup>17</sup>1mas-posizionamento-motore-ricerca.it/, *Statistiche SEMS 2009* [[http://www.1mas-posizionamento-motore-ricerca.it/ita/seo\\_utilizzo\\_motori\\_ricerca\\_italia\\_2009](http://www.1mas-posizionamento-motore-ricerca.it/ita/seo_utilizzo_motori_ricerca_italia_2009)]

spingono al massimo fino alla quarta pagina dei risultati visualizzati, dopodiché cambiano la query o il motore di ricerca utilizzato e iniziano una nuova ricerca. Il 18% degli utenti arriva invece fino alla terza pagina, il 27% fino alla seconda, il 17% solo alla prima pagina, mentre il 6% degli utenti cambia addirittura solo dopo i primi sei risultati visualizzati.

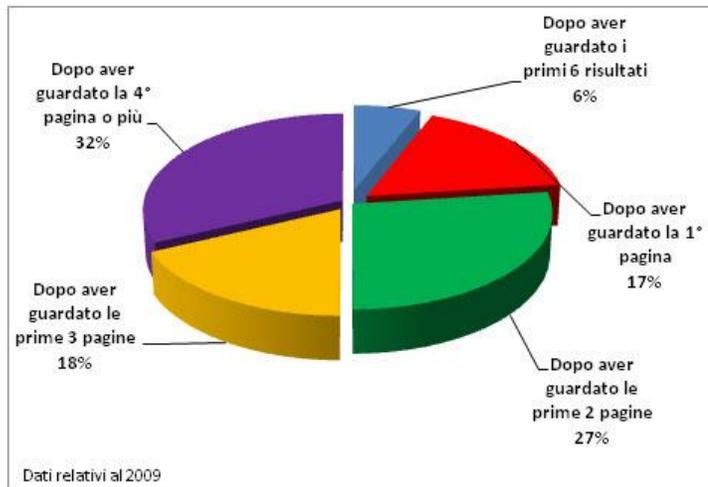


Tabella 2.5 Comportamento degli utenti prima di cambiare keyword o motore dopo una ricerca

Con questi dati, si capisce come sia importante per un sito uscire tra le prima pagine delle SERP, se si vuole che ottenga traffico dai motori di ricerca e visibilità. La SEO, con le proprie strategie, può intervenire su questo aspetto e portare vantaggi a un sito, migliorando il suo posizionamento tra i risultati delle ricerche e aumentando il traffico di visitatori proveniente dai motori di ricerca.

## 2.4 Vantaggi della SEO

La figura 2.2 mostra il layout tipico di una SERP di Google. Sotto il box di ricerca e a destra della pagina appaiono i link sponsorizzati, i cosiddetti *Pay Per Click*, mentre nella parte inferiore della pagina si trovano i risultati di ricerca veri e propri, chiamati *risultati organici*, cioè quelli che possono essere influenzati direttamente dalle attività SEO.



Figura 2.2 Il layout tipico di una SERP di Google

Da notare, che i risultati organici occupano la maggior parte dello spazio disponibile sulla pagina, e questo è particolarmente evidente quando si effettuano ricerche con parole chiavi meno popolari, che hanno pochi link sponsorizzati. Usando parole chiave di nicchia, o parole chiave costituite da più termini, le cosiddette keyword *long tail*, spesso si ha il 100% di risultati organici, senza nessun link sponsorizzato. La figura 2.3 mostra un esempio di una ricerca per le parole chiave *come ottimizzare un sito web*. Come si vede, non esistono link PPC.

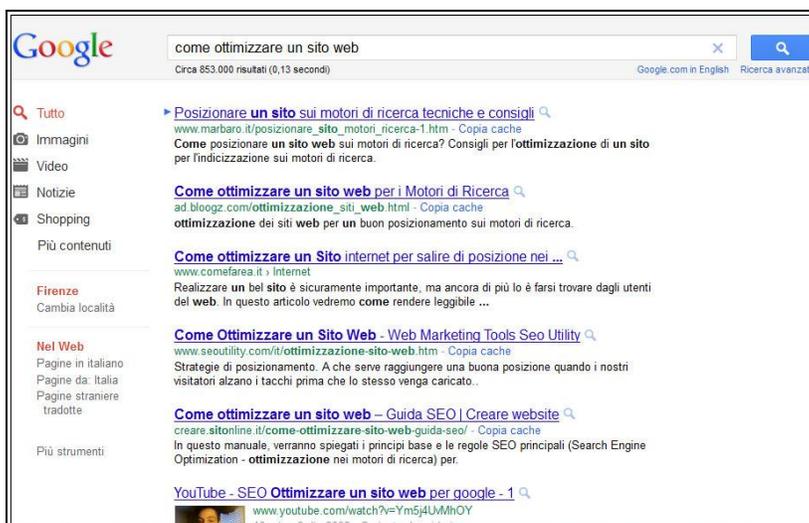


Figura 2.3 Risultati SERP con parole chiave di nicchia

Detto questo, possiamo individuare subito i vantaggi della SEO, che lavora sui risultati organici, e punta su parole chiave di nicchia: aumentare la probabilità di far uscire una pagina tra i primi risultati di una SERP senza la presenza delle campagne PPC e con

parole chiave che possono soddisfare meglio le richieste dell'utente, perché più mirate, aiutando spesso a ottenere maggiori tassi di conversione e trasformando l'utente in un potenziale cliente. Inoltre, affinché la SEO funzioni, i contenuti di un sito devono essere sempre aggiornati, e i motori di ricerca, se vedono che un sito si rinnova costantemente, lo considera utile e ne favorisce il posizionamento. Il tutto, in modo gratuito.

### **Costi di gestione inferiori**

I contatti gratuiti generati dai risultati organici minimizzano o eliminano la necessità di dover effettuare campagne PPC, riducendo notevolmente i costi di marketing e favorendo invece in guadagno con l'aumento dei potenziali clienti.

## **2.5 Sfide della SEO**

La SEO è un'attività valida, ma richiede un impegno a lungo termine, e tra gli aspetti che deve tenere in considerazione, ce ne sono alcuni che non sono legati alla tecnologia ma anzi, dipendono da fattori esterni.

### **Competizione**

Secondo un'indagine condotta da Netcraft<sup>18</sup>, nel Giugno 2011 esistevano poco più di 104 milioni di domini registrati attivi. Sebbene questo sembri un numero sbalorditivo, probabilmente è solo una stima prudenziale, visto che ciascun dominio ha molto spesso altri sotto domini. Quindi, anche se si lavora in un settore di nicchia, è difficile poter realizzare un sito che non abbia concorrenza, perché oggi quasi tutto il mercato si muove sul web. Per fortuna però, non tutti i siti utilizzano la SEO, cosa che avvantaggia gli altri siti nel posizionamento.

### **Assenza di garanzie**

Proprio perché le variabili in gioco sono tante, nessuno può garantire con certezza di riuscire a posizionare un sito in cima ai risultati di Google, Yahoo! o Bing. In ogni caso,

---

<sup>18</sup> Netcraft.com, *2011 Web Server Survey* [<http://news.netcraft.com/archives/2011/06/07/june-2011-web-server-survey.html>]

i vantaggi della SEO sono reali, e se si applica con competenza, ci saranno molte possibilità di raggiungere una buona popolarità e un buon posizionamento.

### **Fluttuazione del posizionamento**

Con i continui sviluppi delle tecnologie di ricerca, i motori utilizzano centinaia di parametri per costruire gli algoritmi di ordinamento dei risultati. A volte basta che anche uno solo di questi parametri cambi, perché un sito precipiti o addirittura sparisca del tutto dai risultati. Sebbene nessuno conosca le regole per un buon posizionamento, ci sono comunque fattori positivi che l'aiutano (su cui punta la SEO) e fattori negativi che lo penalizzano.

### **Impegno a lungo termine**

Ogni sito è diverso, e ognuno di essi ha bisogno di una propria strategia SEO specifica. I fattori in gioco sono così tanti, che sarebbe ingenuo stabilire un tempo esatto per vedere i risultati attesi. Inoltre, anche quando i risultati arrivano, il compito della SEO non finisce, perché bisogna saper mantenere il posizionamento raggiunto. La SEO favorisce lo sviluppo naturale e di lungo periodo di un sito e spesso, una volta ottenuto un buon posizionamento sui motori di ricerca, arrivano anche ulteriori benefici dal traffico proveniente dai link che altri siti iniziano ad utilizzare. Se poi si hanno contenuti adeguati per il target di riferimento, diventa naturale sviluppare traffico in ingresso.

## 3. Il processo di ottimizzazione SEO

Ottimizzare un sito web significa mettere in atto le tecniche che tendono a migliorare la posizione del sito stesso nelle SERP. Il lavoro di implementazione delle tecniche SEO richiede molto tempo e uno studio accurato di diversi importanti fattori. Il processo SEO, come illustrato nella figura 3.1, può essere suddiviso in 6 fasi generali: ricerca, pianificazione strategica, implementazione, monitoraggio, (ri)verifica e manutenzione<sup>19</sup>.

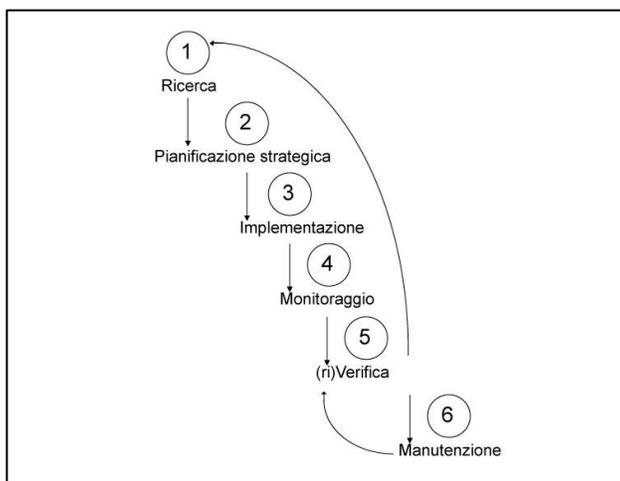


Figura 3.1 Le 6 fasi del processo SEO

### 3.1 Ricerca

La prima fase del processo d'implementazione SEO riguarda la raccolta di informazioni sul settore di riferimento del sito, l'analisi competitiva, l'analisi dello stato attuale del sito e la ricerca delle parole chiave.

#### 3.1.1 Settore di riferimento

Prima di tutto, è necessario conoscere a fondo come il proprio settore viene trattato online. Alcune domande da porsi saranno: quale mercato vogliamo raggiungere? Quali sono i risultati che ci si aspetta di raggiungere? Più si ha una visione iniziale chiara, più le azioni che si andranno a intraprendere saranno focalizzate.

<sup>19</sup> John I. Jerkovic. *SEO. L'arte dell'ottimizzazione per i motori di ricerca*. Milano, Tecniche Nuove, collana Hops Tecnologie, 2010

### **3.1.2 Analisi competitiva**

Dopo aver delineato il quadro generale, è importante capire cosa stanno facendo gli altri. Bisogna cercare di determinare i concorrenti del vostro settore, che posizione occupano nei risultati di ricerca dei motori e come ci sono arrivati, stimare il traffico e cercare di capire quali sono le parole chiave più utilizzate per le ricerche. Più informazioni si raccolgono, meglio è.

### **3.1.3 Stato del sito**

Nel caso in cui si voglia aumentare la visibilità di un sito esistente, bisogna analizzarlo in ogni dettaglio. Sono quelle operazioni che vengono definite in gergo *site clinic*, e sono molto simili a quelle già utilizzate per l'analisi della concorrenza: cambia solo l'oggetto della ricerca. Si deve analizzare il posizionamento del sito nei motori di ricerca, per poi passare ai fattori interni al sito, a livello di pagina e di intero sito, ai fattori esterni che lo influenzano, e infine la piattaforma tecnologica su cui si basa il sito, le modalità di aggiornamento di contenuti, le statistiche di accesso e così via.

### **3.1.4 Parole chiave**

La ricerca delle parole chiave consiste nell'individuare e nel definire i termini più significativi attraverso i quali vogliamo farci trovare dai motori di ricerca. La ricerca delle parole chiave però, non riguarda soltanto i contenuti testuali delle pagine, ma anche i nomi dei domini, i link in ingresso, la loro struttura e molti altri elementi. La domanda di base da porsi è: su quali parole chiave ci si deve concentrare? Si vuole puntare su termini generici o su termini di nicchia?

Per siti che investono su campagne a pagamento, è importante valutare il tasso di conversione. Se questo è basso, può significare costi per l'acquisto di parole chiave superiori ai ritorni ottenuti, cosa che rende inutile lo scopo di una campagna Pay Per Click.

## **3.2 Pianificazione strategica**

La fase di pianificazione strategica deve rispondere ad alcune domande fondamentali che derivano dai risultati ottenuti dalla fase di ricerca. È necessario elaborare diverse strategie che riguardano sia i contenuti, sia la tecnologia.

### **3.2.1 Contenuti**

La strategia sui contenuti deve riguardare tutto il loro ciclo di vita: creazione, modifica, divulgazione e archiviazione. Bisogna considerare come vengono presentati agli utenti, e se il sito deve comprendere eventuali funzionalità aggiuntive come blog, comunicati stampa, feed RSS, testimonianze, file multimediali e così via. Inoltre, bisogna valutare quali contenuti si vogliono rendere indicizzabili dai motori di ricerca e quali no.

### **3.2.2 Sviluppo dei link**

Questo è una delle strategie pilastro della SEO. I contenuti senza link, infatti, non vanno lontano. Bisogna pianificare un'architettura dei link interni per favorire la visibilità delle pagine di un sito agli spider dei motori di ricerca, e avere buoni contenuti per favorire una crescita del traffico in ingresso e uno sviluppo naturale dei backlink che aumentano la popolarità.

### **3.2.3 Piattaforma tecnologica**

È importante considerare la piattaforma tecnologica necessaria per implementare le funzionalità del sito, valutare il supporto dei browser e le prestazioni del sito. Questo aspetto è fondamentale per siti di grandi dimensioni, ma in generale i siti che impiegano molto tempo per caricare le pagine costituiscono una perdita di tempo sia per gli utenti, che possono decidere di uscire dal sito, sia per gli spider, che a loro volta possono decidere di non indicizzare la pagina perché richiede troppo tempo.

## 3.3 Implementazione

La fase di implementazione riguarda le due aree di lavoro centrali: l'ottimizzazione interna e quella esterna del sito. Entrambe queste fasi verranno affrontate nel dettaglio nelle sezioni successive.

### 3.3.1 Ottimizzazione interna

L'ottimizzazione interna fa riferimento sia ad attività sulle singole pagine (*on-page*) che sull'intero sito (*on-site*). In particolare, le attività sulle singole pagine si concentrano nell'ottimizzazione delle parole chiave da inserire nei tag `<title>` e `<meta>`, la verifica redazionale di testi e la definizione dei link. L'ottimizzazione del sito, invece, parte dalla definizione del dominio per passare alla progettazione grafica, alla configurazione del server e all'analisi delle prestazioni.

### 3.3.2 Ottimizzazione esterna

L'ottimizzazione esterna (*off-site*) è altrettanto importante di quella interna. L'obiettivo principale di queste attività è di dare visibilità al sito cercando di favorire la comparsa di link diretti da altri siti. Ogni link infatti, può essere visto come un “voto” per il sito. Più i link sono di qualità e più il sito aumenterà la sua popolarità agli occhi dei motori di ricerca, che in base a questi voti, lo può considerare autorevole e utile per le ricerche degli utenti, migliorando così il suo posizionamento nei risultati delle ricerche.

## 3.4 Monitoraggio

Questa fase riguarda la verifica nel tempo di diversi fattori, come le attività degli spider, la quantità di link da altri siti, il posizionamento sui motori, i tassi di conversione dei link e altro. Tutte queste attività sono collegate tra loro e poggiano sull'analisi dei log<sup>20</sup> del server web. Tra gli strumenti più utilizzati per l'analisi di un sito troviamo Google Webmaster Tools e Google Analytics.

---

<sup>20</sup> Tradotto come *giornale di bordo*, nell'informatica indica la registrazione cronologica delle operazioni man mano che vengono eseguite e il file su cui tali registrazioni sono memorizzate.

### **3.5 Valutazione e manutenzione**

Le fasi di valutazione e manutenzione si occupano di analizzare i dati raccolti durante il monitoraggio e risolvere eventuali problemi venuti alla luce durante le fasi di (ri)valutazione. Questa fase prevede la verifica periodica delle metriche di base, cercando di capire cosa funziona e cosa invece bisogna correggere. Come mostrato nella figura 2.5, i problemi più semplici possono essere risolti nella fase di manutenzione, mentre quelli più complessi possono richiedere di dover tornare alla fase di ricerca.

## 4. Strategia per le parole chiave: analisi e ricerca

Le parole chiave sono i termini o le frasi utilizzate dagli utenti quando cercano qualcosa dai motori di ricerca. È importante quindi definire una strategia chiara e coerente, scegliendo le giuste combinazioni di parole chiave da utilizzare all'interno di un sito.

### 4.1 La scelta delle parole chiavi corrette

In generale, si possono scegliere due tipologie di parole chiave: generiche e specifiche. Chiaramente, più una parola è generica, più è soggetta a forte concorrenza. Puntare su termini generici richiede quindi molti sforzi, e un buon posizionamento nei risultati è molto difficile da raggiungere.

Termini più specifici invece, oltre che a essere più facili da ottimizzare, consentono una precisione maggiore nella ricerca e quindi portano un'utenza molto più targhettizzata e presumibilmente più soddisfatta (cosa che aumenta il tasso di conversione).

#### 4.1.1 La Teoria delle *Long Tail*

Il termine *long tail* (coda lunga) è un termine coniato da Chris Anderson nel 2004<sup>21</sup> per indicare la forma della distribuzione statistica dell'utilizzo di parole chiave specifiche nelle ricerche, in funzione del numero di parole che costituiscono le ricerche stesse<sup>22</sup>. La figura 4.1 mostra la forma tipica di questa distribuzione, dove si nota che le parole chiave specifiche che utilizzano più termini, sono quelle soggette a minore competizione. Sebbene queste parole chiave non abbiano i volumi di utilizzo tipici delle parole generiche, godono di altri vantaggi, tra cui una minore competitività e maggiori tassi di conversione.

---

<sup>21</sup> Wikipedia, voce *Long Tail* [[http://en.wikipedia.org/wiki/Long\\_Tail](http://en.wikipedia.org/wiki/Long_Tail)]

<sup>22</sup> John I. Jerkovic. *SEO. L'arte dell'ottimizzazione per i motori di ricerca*. Milano, Tecniche Nuove, collana Hops Technologie, 2010

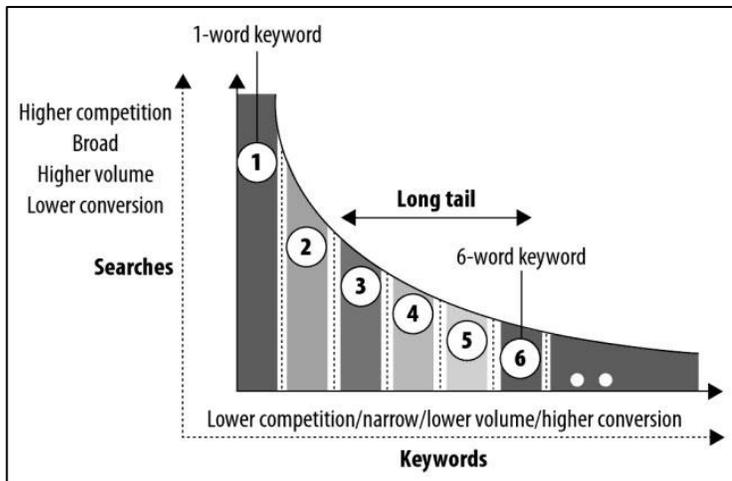


Figura 4.1 Parole chiavi e long tail

Un tasso di conversione più alto nelle parole chiave *long tail* è una conseguenza naturale del fatto che chi le usa nelle ricerche, troverà risultati molto più concreti. In ogni caso, utilizzare una strategia *long tail* non implica necessariamente l'utilizzo di parole chiave costituite da molti termini. Bisogna cercare un equilibrio tra l'effettivo volume di utilizzo dei termini e la lunghezza delle frasi di ricerca su cui si vuole puntare.

#### 4.1.2 Parole chiave e lessico

Una parola chiave può essere un qualsiasi termine. Può essere un nome (es. *sito*, *ottimizzazione*, *Jakob Nielsen*), un aggettivo (es. *usabile*, *accessibile*), un avverbio (es. *velocemente*) un verbo (es. *posizionare*, *ottimizzare*), oppure un insieme di parole (es. *motore di ricerca*, *sito web*) o una frase (es. *posizionare un sito web velocemente*). Alcuni dei concetti da prendere in considerazione durante lo studio delle parole chiave, utilizzati dai motori durante l'indicizzazione, sono le parole derivate, i modificatori di significato e il Latent Semantic Indexing.

#### L'importanza delle parole derivate

Le parole derivate sono variazioni di un termine a partire da una stessa radice. Questo concetto è importante, poiché l'utilizzo di parole derivate nelle pagine web può aiutare a incrementarne la rilevanza rispetto ai termini di ricerca, e quindi a ottenere traffico proveniente dalle SERP. Google infatti, nel calcolare il posizionamento e la rilevanza

delle pagine web, non considera solo i termini così come vengono inseriti, ma anche le loro derivazioni, attraverso un sofisticato processo di stemming durante fase di indicizzazione (vedi p. 9).

### **Modificatori di significato**

I modificatori di significato delle parole chiave, sono termini che compaiono immediatamente prima o dopo una parola chiave. Sono importanti perché anche loro possono rendere più efficaci le strategie per le parole chiave *long tail*.

Tra i tipi di modificatori troviamo quelli generici e di nicchia. I modificatori generici possono essere utilizzati in qualsiasi sito e sono normalmente termini che compaiono spesso nelle ricerche effettuate dagli utenti. Quelli di nicchia invece sono termini specifici di un dato settore. Partendo dalla radice delle parole base, si possono escogitare diverse variazioni di termini e di parole correlate. In questo modo, i motori indicizzeranno il sito da vari punti di vista e gli assegneranno una rilevanza tanto maggiore quanto più saranno correlati tra loro i termini trovati.

### **Latent Semantic Indexing (LSI)**

La Latent Semantic Indexing (Indicizzazione Semantica Latente), conosciuta anche come LSA (Latent Semantic Analysis) è un metodo di indicizzazione e analisi che utilizza una tecnica matematica per identificare schemi ricorrenti nelle relazioni tra i termini e i concetti contenuti in un sito<sup>23</sup>. La capacità di questo algoritmo, utilizzato dai motori di ricerca per migliorare la qualità dei risultati, sta quindi nell'estrarre il contenuto concettuale dal corpo del testo e stabilire associazioni tra quei termini che ricorrono in contesti simili. Ad esempio, in un sito che parla di *web design*, si presume che le parole chiave più prevedibili siano relative alla grafica, all'usabilità, all'accessibilità ecc. visto che tutti questi termini sono rilevanti per il tema del sito. La LSI costituisce anche un valido aiuto per superare le difficoltà che hanno i motori di ricerca nel perfezionamento dei risultati, specialmente per la disambiguazione di una parola, quando cioè ci si trova di fronte a parole diverse che hanno significati simili (sinonimi) e a parole che invece possiedono più di un significato (polisemiche).

---

<sup>23</sup>Wikipedia, voce *Latent Semantic Indexing* [[http://en.wikipedia.org/wiki/Latent\\_semantic\\_indexing](http://en.wikipedia.org/wiki/Latent_semantic_indexing)]

Dal punto di vista SEO, invece, la LSA costituisce un'occasione per porre l'attenzione nella fase di redazione dei contenuti delle pagine da ottimizzare: utilizzare non solo le parole chiave principali, ma anche parole strettamente correlate all'argomento trattato, se non si vuole correre il rischio di essere considerati poco pertinenti dai motori di ricerca per un determinato argomento, ed essere quindi penalizzati nel posizionamento.

## 4.2 Gli strumenti utili per le parole chiave

Durante la fase di ricerca delle parole chiave, è importante avere le idee molto chiare e conoscere a fondo il settore su cui si deve lavorare. Gli strumenti che suggeriscono le parole chiave da utilizzare sono tanti, gratuiti e a pagamento. Qui, verranno illustrati alcuni degli strumenti di Google più utilizzati.

### Google AdWords Keyword Tool

È forse il più diffuso strumento gratuito per la ricerca delle parole chiave e uno dei migliori disponibili. Le parole chiave possono essere cercate in due modi: inserendo un elenco di termini, oppure inserendo la URL di un sito. Grazie ai molti filtri disponibili, si può effettuare una ricerca personalizzata e più approfondita. Come mostrato in figura 4.2, per ogni risultato della ricerca il Keyword Tool indica il livello di concorrenza tra inserzionisti per quella determinata parola chiave o frase, il volume di ricerca mensile globale e il volume di ricerca mensile locale.

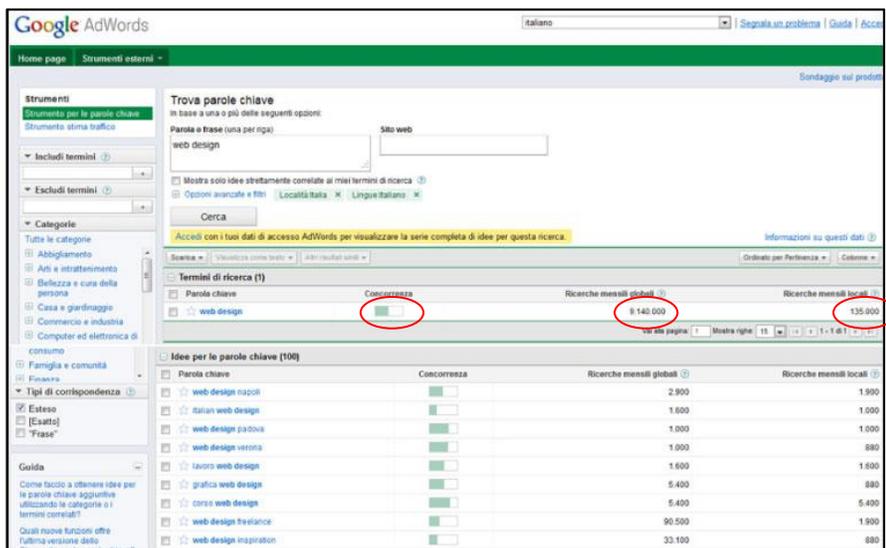


Figura 4.2 Google AdWords Keyword Tool

Per essere più sicuri sul volume effettivo di ricerca mensile globale di una parola chiave, si può fare un riscontro con il numero di pagine effettivamente indicizzata da Google per quella keyword, come mostrato nell'esempio in figura 4.3.



Figura 4.3 Visualizzazione numero di pagine concorrenti

## Google Trends

Google Trends, come si vede in figura 4.4, è uno strumento che visualizza l'andamento nel tempo dell'utilizzo di un dato termine da parte degli utenti nelle ricerche. Google Trends consente di inserire nel box di testo fino a un massimo di 5 keywords, separate dalla virgola. Lo strumento visualizza come risultato due tipologie di grafico: nel primo viene raffigurato il volume di ricerca con cui il termine è stato ricercato in Google, nel secondo invece viene indicato quante volte i news media hanno trattato l'argomento. È possibile inoltre impostare anche un filtro per regione geografica specifica, e per un dato intervallo di tempo.



Figura 4.4 Google Trends

## Google Insights for Search

Google Insight for Search visualizza un grafico che riporta il numero di ricerche effettuate per un determinato termine rispetto a tutte le ricerche effettuate in Google in

un determinato arco temporale. Anche qui, è possibile selezionare la regione di interesse per visualizzare il volume di ricerca delle keywords, e analizzare i termini che sono strettamente correlati con i termini cercati e che hanno un elevato tasso di ricerca da parte degli utenti Google. Nella stessa sezione, è possibile inoltre visualizzare anche le ricerche emergenti, che rappresentano una serie di frasi chiave che hanno fatto registrare un notevole incremento di volume in un determinato periodo di tempo, rispetto al periodo di tempo precedente. Nella figura 4.5 è riportato un esempio.

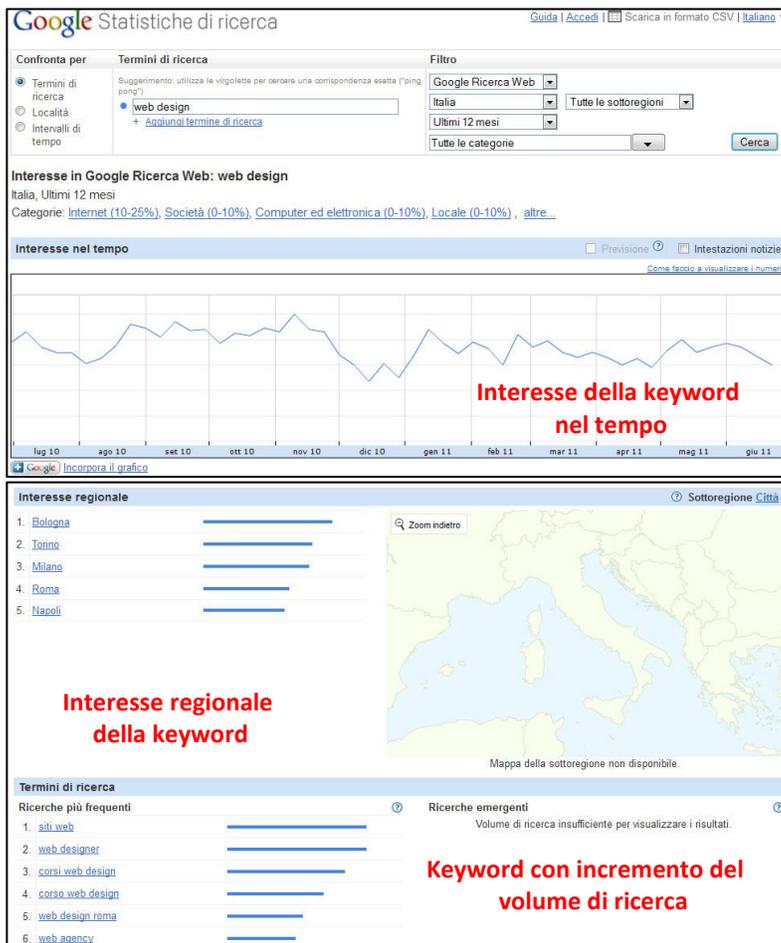


Figura 4.5 Google Insight for Search

## 5. Ottimizzazione interna on-page

I motori di ricerca come Google, Yahoo! e Bing, nelle loro guide per i webmaster, illustrano i fattori on-page più importanti che sono in grado di influenzare il posizionamento di una pagina web nei risultati delle ricerche, e vengono analizzati di seguito.

### 5.1 Head Tags

Gli head tags rappresentano quei tag<sup>24</sup> che si trovano nella parte di intestazione di un documento HTML, che viene definita appunto <head>. Il loro corretto utilizzo è fondamentale, non solo perché alcuni di questi tag vengono mostrati agli utenti nei risultati delle SERP, ma perché costituiscono importantissimi fattori di indicizzazione per i motori di ricerca.

#### 5.1.1 Tag Title

Il tag <title> rappresenta il fattore interno della pagina più importante in assoluto. I motori di ricerca, utilizzano il testo in questo tag anche come titolo da far comparire nei risultati delle ricerche, come mostrato nella figura 5.1. Da notare poi, che nelle SERP vengono evidenziate in grassetto le parole che rappresentano esattamente le parole digitate nel box di ricerca.



Figura 5.1 visualizzazione dei tag *title* nei risultati della ricerca

<sup>24</sup> Indica un elemento inserito nel testo di un documento HTML per indicarne la struttura o la formattazione.

Bisogna quindi scegliere un titolo che contenga le parole chiave più importanti, ma allo stesso tempo, non deve essere solo un accostamento di parole chiave, ma deve essere pertinente, descrittivo e accattivante, dando all'utente la sensazione di aver trovato quello che stava cercando e indurlo così a entrare nel sito. Per queste ragioni, è molto importante dedicarsi con cura alla scelta del titolo, rendendolo sempre unico e originale. Inoltre, è buona norma utilizzare titoli brevi, perché i motori di ricerca non visualizzano nei risultati oltre i sessanta, settanta caratteri. In questo modo, si evita che un titolo troppo lungo venga troncato e non venga visualizzato. Quindi, è meglio mettere all'inizio del titolo le parole chiave più importanti, in modo che siano subito visibili dall'utente e dal motore di ricerca.

### 5.1.2 Meta tag Description

Il tag <meta> *description*, come si vede in figura 5.2, è quel breve testo che appare sotto al titolo nell'elenco dei risultati e serve a dare una breve descrizione del contenuto della pagina.



Figura 5.2 Visualizzazione del meta tag *description* nei risultati della ricerca

Benché questo meta tag non influisca sull'indicizzazione, si rivela molto utile. I motori di ricerca infatti, utilizzano il testo all'interno di questo tag per fornire informazioni sulla pagina che appare nei risultati delle ricerche, orientando l'utente nella scelta e aumentando così le probabilità di click a un sito. Una buona *description* deve quindi contenere una sintesi dei contenuti presenti nella pagina web, essere accattivante e diversa per ogni pagina. È consigliabile che la lunghezza della descrizione non superi i

160 caratteri, per evitare che il testo non venga mostrato per intero. I motori di ricerca, inoltre, tendono a visualizzare uno *snippet*, così viene chiamato il breve testo descrittivo che appare nelle SERP, che sia coerente con quanto cercato dall'utente. Per cui, se la *description* manca, o in essa non sono inserite le parole chiave che corrispondono alla ricerca effettuata, i motori vanno a cercare la prima frase a caso nel corpo della pagina che contenga le keywords cercate dall'utente.

### 5.1.3 Meta tag Keyword

Il tag `<meta> keyword`, si utilizza per facilitare lo spider nell'indicizzazione della pagina. Come si vede nella figura 5.3, all'interno del tag si inseriscono un elenco di parole chiave contenute nella pagina, separate da virgola. Una volta questo meta tag costituiva uno dei principali fattori di indicizzazione. A causa del suo utilizzo smisurato e poco onesto, nel tempo i motori di ricerca ne hanno diminuito quasi del tutto l'utilità, tanto che alcuni di essi, come Google, addirittura lo ignorano. Per questo motivo, è inutile inserire nel tag una lunga serie di chiavi, ma se si vuole utilizzare, si consiglia di inserirvi esclusivamente le chiavi che vengono usate per ottimizzare il testo della pagina, in genere non più di cinque.

```
<meta name= "keywords" content= "seo, strategie seo, search engine  
optimisation, ottimizzazione siti web" />
```

Figura 5.3 Esempio di meta tag keyword

### 5.1.4 Meta tag Robots

Il tag `<meta> robots` rientra nel protocollo Robots Exclusion Protocol (REP), ideato nel 1993 come meccanismo per impedire, o consentire, esplicitamente che specifiche pagine web vengano visitate dagli spider, e quindi indicizzate. L'utilizzo del REP si rivela molto importante, ad esempio, per nascondere pagine con contenuti che non si vuole far apparire nei risultati, oppure pagine con contenuti duplicati, uno dei problemi più frequenti per chi si occupa di SEO. Questo tag non fornisce nessuna informazione

relativa ai contenuti di una pagina, ma si limita a fornire direttive allo spider sulla singola pagina che sta visitando.

Le direttive che possono essere impartite sono di due tipi e hanno lo scopo di:

- richiedere l'inclusione (o l'esclusione) dei contenuti della pagina stessa nell'archivio del motore di ricerca;
- richiedere di seguire (o ignorare) tutti i link interni alla pagina che consentono il collegamento ad altre pagine web.

La figura 5.4 mostra la struttura del meta tag *robots*, e di seguito sono indicati i comandi utilizzati per impartire queste direttive agli spider:

- **Index**: comunica allo spider di archiviare la pagina nel database
- **Noindex**: comunica allo spider di non archiviare la pagina nel database
- **Follow**: comunica allo spider di seguire i link presenti nella pagina
- **Nofollow**: comunica allo spider di non seguire i link nella pagina

```
<meta name= "robots" content="index, follow" />
```

Figura 5.4 Esempio di meta tag robots

Benché i motori di ricerca abbiano deciso di supportare in modo uniforme e coerente il protocollo REP, non è obbligatorio conformarsi alle sue regole, quindi alcuni motori di ricerca non lo fanno. In questo caso, gli spider non hanno nessuna direttiva da seguire perché non interpretano i comandi del meta tag robots e quindi, archiviano la pagina nel database e seguono i link contenuti nella pagina.

## 5.2 Body Tags

I tag che si trovano nel corpo del testo sono molto importanti, perché consentono di ottimizzare l'aspetto fondamentale di una pagina web: i contenuti. Di seguito vengono analizzati tutti i tag all'interno del corpo di una pagina che aiutano l'indicizzazione.

### 5.2.1 Heading Tag

Gli *heading tags*, o tag di intestazione, sono utili per fornire una struttura logica al contenuto di una pagina, attraverso una gerarchizzazione delle sue parti comprensibile agli spider. Perciò, imparare a utilizzare correttamente questi tag all'interno di un blocco di testo, è importante per aiutare l'indicizzazione, perché segnalano quali parole sono più importanti rispetto ad altre all'interno di un testo. L'utente è in grado di individuarle in base alla dimensione dei caratteri, più grande rispetto al resto del testo, a seconda del valore che è stato attribuito al tag. Gli heading tags definiscono questa gerarchia attraverso una scala che va da 1 a 6:

- `<h1>`: è il tag che racchiude il contenuto di maggiore importanza. In genere corrisponde al titolo della pagina e per questo va utilizzato solo una volta all'interno della stessa pagina;
- `<h2>`: viene generalmente utilizzato per indicare il sottotitolo di una pagina e di conseguenza appare con una dimensione un po' più piccola rispetto a quella inserita in `<h1>`;
- `<h3>`, `<h4>`, `<h5>`, `<h6>`: hanno un'importanza minore rispetto agli altri due tag, per cui se ne consiglia l'utilizzo solo se necessario.

È importante tener presente che questi tag non hanno uno scopo grafico, ma solo strutturale. Bisogna quindi porre attenzione a come si utilizzano, ricordando che in essi vogliamo inserire una parte di testo che vogliamo evidenziare sia all'utente, sia agli spider. Dobbiamo perciò scrivere una descrizione concisa e chiara, ma che contenga anche le parole chiave su cui vogliamo puntare.

## 5.2.2 Anchor text

L'*anchor text* corrisponde a quella parte di testo della pagina cliccabile da parte dell'utente, che comunemente è nota con il termine: *link*. Il corretto utilizzo di questo tag, e una sua giusta architettura, non solo facilita la navigazione dell'utente all'interno di un sito, ma costituisce uno dei fattori più importanti di indicizzazione. Gli algoritmi dei motori di ricerca infatti, considerano molto importanti i testi contenuti in un link, per capire gli argomenti trattati in quella pagina e per evidenziare queste pagine linkate agli spider. Link troppo lunghi o generici sono penalizzanti. Al contrario, un testo descrittivo e conciso, che contenga un paio di parole chiave giuste, e che sia facilmente riconoscibile dall'utente rispetto al resto del testo, è sicuramente un fattore positivo per l'indicizzazione. Quindi, più un anchor text è chiaro e semplice, più sarà comprensibile e utile per l'utente e i motori di ricerca. La figura 5.5 mostra un esempio di anchor text ben strutturato, mentre la figura 5.6 mostra come questo codice verrebbe visualizzato da un browser:

```
<a href="http://www.strategieseo.com/articoli/le-strategie-seo-per-un-buon-posizionamento-nei-motori-di-ricerca.htm" title="Strategie SEO per posizionarsi nei motori di ricerca"> Strategie SEO </a>
```

Figura 5.5 Un esempio corretto di anchor text lato SEO

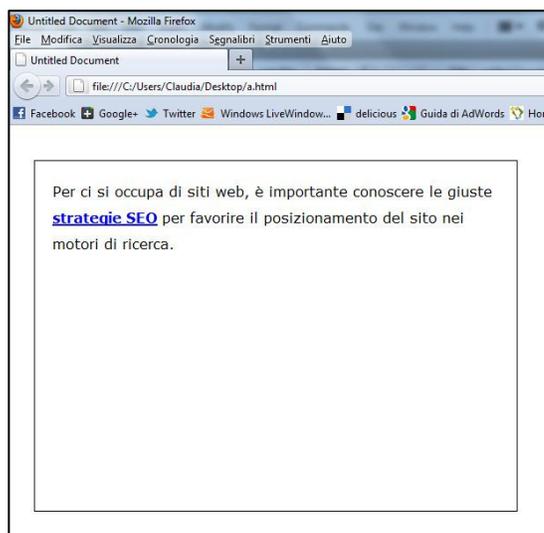


Figura 5.5 Un esempio corretto di anchor text visualizzato sul browser

### 5.2.3 Le immagini e il tag alt

Le immagini costituiscono una componente spesso sottovalutata ai fini dell'ottimizzazione, ma in realtà si rivelano una risorsa molto utile per aiutare il posizionamento. Questo è particolarmente evidente negli ultimi tempi, grazie agli aggiornamenti sempre più sofisticati degli algoritmi di Google. Infatti, le SERP stanno diventando sempre più complete, visualizzando risultati non solo di le pagine web, ma anche di immagini e video, come nell'esempio mostrato in figura 5.7.



Figura 5.7 Esempio di SERP Google con immagini e video

Per ottimizzare le immagini, occorre prestare attenzione ad alcuni elementi fondamentali. Il nome del file deve essere chiaro e descrittivo, evitando quindi nomi privi di informazione come *DSCN1234.jpg*. Altrettanto vale per il testo che inseriamo negli attributi `title` e `alt` del tag, considerato molto utile dai motori di ricerca, perché permette di comprendere il contenuto dell'immagine e di utilizzarlo ai fini dell'indicizzazione. La figura 5.8 mostra il codice utilizzato per ottimizzare un'immagine. Il `title` corrisponde al titolo, e viene visualizzato sul browser quando l'utente passa il cursore del mouse sopra l'immagine.

```

```

Figura 5.8 Ottimizzazione di una immagine lato SEO

L'attributo `alt` invece, come si vede in figura 5.9, si utilizza per dare una descrizione dell'immagine, in modo che se questa non dovesse essere visibile all'utente per problemi di server o altro, si fornisce un'informazione *alternativa*, con del testo che ne descrive il contenuto, aiutando così anche l'accessibilità.

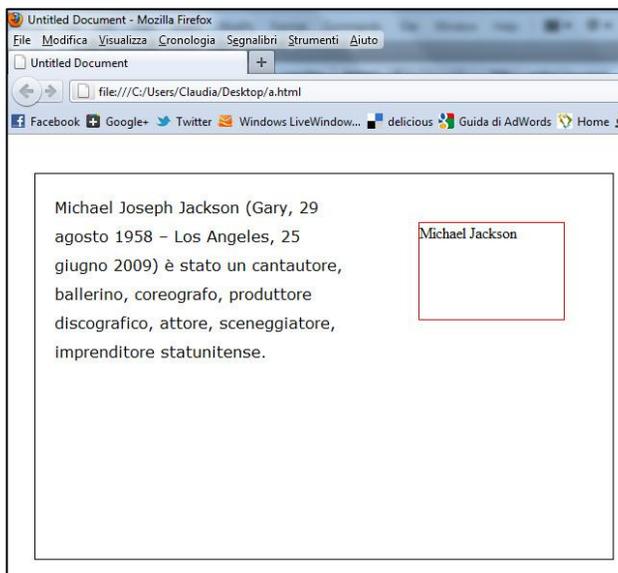


Figura 5.9 L'attributo `alt` visualizzato sul browser

Per una corretta ottimizzazione, è importante che il `title` e l'`alt` siano sempre diversi per ogni immagine, anche perché, nel caso in cui l'immagine sia linkabile, il testo contenuto nell'attributo `alt` viene interpretato dagli spider come se fosse il testo di un anchor text, facilitando i motori di ricerca nell'interpretazione del contenuto della pagina. E' sempre bene utilizzare formati per le immagini che siano supportati dai browser, come ad esempio `jpg`, `png` o `gif`, e inoltre, è buona norma inserire le immagini all'interno di un'unica cartella, in modo da semplificare allo spider il percorso da seguire per poterle raggiungere. Infine, un utile suggerimento, è quello di aggiungere una eventuale Sitemap delle immagini, in modo che gli spider abbiano informazioni più precise e dettagliate una volta che trovano le immagini.

## 5.3 Densità delle parole chiave

La densità delle parole chiave costituisce uno dei parametri più importanti per migliorare il posizionamento. Essa rappresenta la percentuale di volte che una determinata parola compare all'interno di una pagina web, rispetto al numero totale delle parole nel testo<sup>25</sup>. Per questo, viene utilizzata dai motori di ricerca come parametro per stabilire la pertinenza di una pagina relativa a una determinata keyword. Infatti, attraverso la stima della densità di una parola chiave nel testo, i motori sono in grado di capire i temi trattati all'interno di una determinata pagina e, di conseguenza, di individuare i termini di ricerca con i quali indicizzare le pagine. In genere, la densità tipica è compresa in valori che vanno dallo **0,2%** al **4%**. Superata questa soglia, si può verificare il cosiddetto *keyword stuffing*, ovvero una presenza eccessiva di parole chiave nel testo e nei meta tag della pagina, ottenuta mediante tecniche SEO scorrette<sup>26</sup>. I motori di ricerca, quando rilevano una quantità eccessiva di parole chiave in una pagina, la valutano come spam. In questo caso, gli spider apportano penalizzazioni al sito web o, nella peggiore delle ipotesi, la sua esclusione dagli indici dei motori di ricerca. Per poter calcolare la keyword density, basta utilizzare uno dei tanti strumenti gratuiti messi a disposizione sul web, dove inserendo l'URL del sito da analizzare, verranno calcolate le keyword con la frequenza di utilizzo più alta e i valori della loro densità. Un esempio dell'utilizzo di questo strumento è mostrato in figura 5.10, ed è tratto dal sito [www.tuttowebmaster.it](http://www.tuttowebmaster.it).

**Keyword density**

Una delle strategie più semplici per scalare il posizionamento nei motori di ricerca è quella di ottimizzare la keyword density delle pagine del sito web.

Nella tabella sottostante sono visualizzate le 10 parole usate con maggior frequenza sulla pagina e i valori della loro densità, cioè, il valore in percentuale confrontato con il totale delle parole usate. Questa informazione è utile per determinare se e come ampliare eventualmente il numero di parole chiavi mirate.

Pagina analizzata: <http://www.toscanaper.me>  
Totale parole: 510  
Parole ottimizzate, con densità tra il 2% e 7% del totale della pagina: 4 OK

**Risultati generali**  
Keyword stuffing: OK - non sussiste  
Parole ottimizzate nel tag Title: 1 OK  
Parole ottimizzate nel tag Description: 1 OK  
Parole ottimizzate nel tag Keywords: 1 OK  
Rating: [★★★★★](#)

**Annunci Google**  
[Posizionamento Site](#)  
[Keyword](#)  
[Primi Motori Di Ricerca](#)  
[Seo](#)

Parola	Quantità	Densità	Nel tag Title	Description	Keywords
pisa	14	2.75%	No	No	No
luglio	14	2.75%	No	No	No
toscana	11	2.16%	SI	SI	SI
apt	11	2.16%	No	No	No
estate	10	1.96%	No	No	No
lucca	8	1.57%	No	No	No
eventi	8	1.57%	No	SI	SI
me	8	1.57%	SI	No	No
toscanaper	8	1.57%	SI	No	No
vivere	8	1.57%	SI	SI	No

Figura 5.10 Calcolo della keyword density

<sup>25</sup> Wikipedia, voce Keyword density [[http://en.wikipedia.org/wiki/Keyword\\_density](http://en.wikipedia.org/wiki/Keyword_density)]

<sup>26</sup> Wikipedia, voce Keyword stuffing [[http://en.wikipedia.org/wiki/Keyword\\_stuffing](http://en.wikipedia.org/wiki/Keyword_stuffing)]

## 5.4 Prominenza delle parole chiave

La prominenza indica la posizione delle parole chiave rispetto all'inizio della pagina HTML<sup>27</sup>. Tra le zone della pagina su cui occorre prestare particolarmente attenzione alla prominenza, troviamo il tag <title>, le intestazioni e i paragrafi del testo che appaiono in cima alla pagina. Infatti, maggiore è la vicinanza con la zona presa in considerazione, maggiore sarà la visibilità delle keyword all'utente, e di conseguenza, maggiore sarà il peso che i motori di ricerca assegneranno a queste keywords. Inoltre, la prominenza aumenta d'importanza con l'aumento delle dimensioni delle pagine. Per cui, se si hanno pagine con molto testo, conviene cercare di utilizzare le parole chiave più importanti all'inizio della pagina. Infine, è bene ridurre al minimo l'uso del codice Javascript e le formattazioni CSS inframmezzate al testo, isolandole in file separati, in modo da migliorare il lavoro degli spider.

## 5.5 Prossimità delle parole chiave

La prossimità indica quanto vicine sono nel testo della pagina due o più parole chiave<sup>28</sup>. Se una o più keyword si ripetono all'interno del testo, e queste ripetizioni sono vicine tra loro, si viene a creare ciò che viene definito un "punto di concentrazione" delle parole chiave. Questa concentrazione, può aumentare la rilevanza delle keyword all'interno del paragrafo in cui si trovano e, di conseguenza, aumentare la rilevanza dell'intera pagina. Google, ad esempio, utilizza la prossimità come uno dei parametri per stabilire la rilevanza di una determinata pagina.<sup>29</sup> Se possibile, però, è meglio evitare di separare le parole chiave con le cosiddette *stop words* (vedi. p. 9), che abbassano il livello di prossimità e che i motori di ricerca in genere non indicizzano. È altrettanto vero però, che la prossimità delle keyword non deve compromettere la leggibilità di un testo.

---

<sup>27</sup> John I. Jerkovic. *SEO. L'arte dell'ottimizzazione per i motori di ricerca*. Milano, Tecniche Nuove, collana Hops Technologie, 2010

<sup>28</sup> John I. Jerkovic. *SEO. L'arte dell'ottimizzazione per i motori di ricerca*. Milano, Tecniche Nuove, collana Hops Technologie, 2010

<sup>29</sup> Abhishek Das, Ankit Jain, "Indexing the World Wide Web: The Journey So Far", *Next Generation Search Engines: Advanced Models for Information Retrieval*, 2011

[[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/en//pubs/archive/37043.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en//pubs/archive/37043.pdf)]

Bisogna quindi stare attenti a non creare periodi pieni di keywords, sgrammaticati o artificiosi, che non solo comportano la perdita di leggibilità di un testo, ma aumentano le probabilità di penalizzazione nel posizionamento, poiché i motori di ricerca potrebbero interpretarlo come spam.

## 6. Ottimizzazione interna on-site

Dopo aver analizzato i fattori che riescono a influenzare il posizionamento sulle singole pagine, adesso passiamo a quelli che sono validi per l'intero sito.

### 6.1 Le URL

L'URL (Uniform Resource Locator) rappresenta in modo univoco l'indirizzo di una pagina web. Proprio per il suo aspetto identificativo, è importante perché permette di accedere a una risorsa sul web. Focalizzarsi sulla struttura delle URL di un sito, è quindi uno tra i primi aspetti che un esperto SEO deve tenere in considerazione durante la sua pianificazione strategica. Ricordiamo che la URL è il terzo elemento che appare nei risultati delle ricerche, insieme al title e alla description. Quindi, il primo passo, sarebbe quello di contenere la (o le) parole chiave su cui vogliamo puntare maggiormente nel nome del dominio. Questo rappresenta un notevole vantaggio, perché aiuta il posizionamento e spinge l'utente a visitare il sito. Attualmente, riuscire a contenere nel nome del dominio una o più parole chiave comuni è molto difficile, poiché quasi tutte sono ormai già occupate. Possiamo però lavorare sulla struttura delle altre URL del sito, in modo che gli spider riescano a indicizzare meglio tutte le pagine, arrivando a ottimi risultati. Per una corretta ottimizzazione delle URL, bisogna che queste siano brevi e scritte in modo chiaro, che contengano le parole chiave principali su cui si vuole puntare per quella pagina, e possibilmente utilizzare il trattino per separare le parole. Usando il trattino infatti, i motori di ricerca riescono a interpretare le parole come separate tra loro, facilitando l'indicizzazione (non il posizionamento). L'underscore “\_” invece, non viene considerato come un separatore, e le parole vengono interpretate come una sola. Ad esempio, *strategie-seo-per-motori-di-ricerca.html* verrà interpretato come sette parole distinte, mentre *strategie\_seo\_per\_motori\_di\_ricerca.html* sarà visto dai motori di ricerca come un'unica parola, indicizzando in modo poco preciso la pagina.

## URL dinamiche

Una URL dinamica, come mostrato in figura 6.1, è una pagina nel cui indirizzo web sono contenute delle variabili.

```
http://www.strategieseo.it/articoli.php?id=30
```

Figura 6.1 Esempio di URL dinamica

In questo caso, la variabile è chiamata *id* e il suo valore è 30. In genere, questo avviene quando i siti si appoggiano a un *database* che contiene tutte le informazioni, e, a seconda delle variabili che gli utenti richiamano, mostrano un risultato. Fino a qualche tempo fa, i motori di ricerca non riuscivano a indicizzare questo tipo di pagine, perché le combinazioni possibili di URL apparivano infinite. Già dal 2008 però, Google, ha dichiarato di aver risolto il problema<sup>30</sup>, mentre nel 2010 Yahoo! e Bing, ancora preferivano l'utilizzo delle URL statiche<sup>31</sup>. In ogni caso, è sempre bene utilizzare il più possibile URL statiche, perché alcuni motori hanno ancora diversi problemi al riguardo, perché si possono inserire parole chiave nell'URL e perché è più facile, per un utente, da memorizzare. Quando è impossibile fare a meno delle pagine dinamiche, si possono ottimizzare le URL con il cosiddetto URL Rewriting. Attraverso alcuni file di istruzione, si cambia il formato di una URL, indicando al nostro server come deve interpretarla. Tutto dipende dal tipo di server che ospita il sito, se Linux o Windows. Ad esempio, dall'URL dinamica:

```
strategieseo.com/articoli.php?id=3titolo=on-page-seo
```

si può ottenere:

```
strategieseo.com/articoli/on-page-seo.html
```

## URL canonici

Gli URL canonici permettono di evitare che i motori di ricerca interpretino una pagina come contenuto duplicato. I motori di ricerca infatti, tendono a non presentare contenuti duplicati in una stessa ricerca, e lo fanno attraverso penalizzazioni su documenti e siti.

---

<sup>30</sup> Google Webmaster Central Blog [<http://googlewebmastercentral.blogspot.com/2008/09/dynamic-urls-vs-static-urls.html>]

<sup>31</sup> Yahoo! Search Help [<http://help.yahoo.com/l/us/yahoo/search/siteexplorer/dynamic-01.html>]

La causa dei contenuti duplicati spesso non è volontaria. A volte capita, per particolari esigenze tecniche, di creare pagine diverse con contenuti simili. Per esempio, si può trattare di un sito di e-commerce, che contiene più pagine con gli stessi prodotti catalogati in ordine differente, oppure di contenuti archiviati o versioni stampabili di una pagina. Grazie a Google, sono state introdotte le URL canoniche, con le quali si può specificare una pagina canonica, preferenziale, per i motori di ricerca, aggiungendo un elemento `<link>` con l'attributo `rel="canonical"` alla sezione `<head>` in tutte le versioni *non* canoniche della pagina. Aggiungendo questo link e questo attributo, si suggerisce a Google la pagina più utile tra tutte le pagine con contenuti identici, indicando di darle la priorità nei risultati di ricerca. La figura 6.2 mostra un esempio di sintassi per specificare una URL canonica:

```
<link rel="canonical" href="http://www.strategieseo.com  
/articoli.php?id=30"/>
```

Figura 6.2 Esempio di URL canonica

## 6.2 Struttura del sito

Una delle caratteristiche fondamentali per un sito web è che sia facilmente navigabile, in modo da aiutare l'utente a trovare rapidamente ciò che sta cercando e fruire meglio le informazioni presenti. La navigabilità di un sito però, risulta estremamente importante anche per i motori di ricerca, per capire quali pagine sono più importanti rispetto ad altre, e aiutare l'indicizzazione. Per questo, occorre prestare particolare attenzione alla strutturazione di un sito web. Una delle strutture che offre migliori risultati è quella denominata ad albero, perché, navigando il sito in profondità, permette all'utente e ai motori di raggiungere in modo sempre più dettagliato e specifico le informazioni. La figura 6.3 illustra un ipotetico schema della struttura di un sito che parla di SEO.

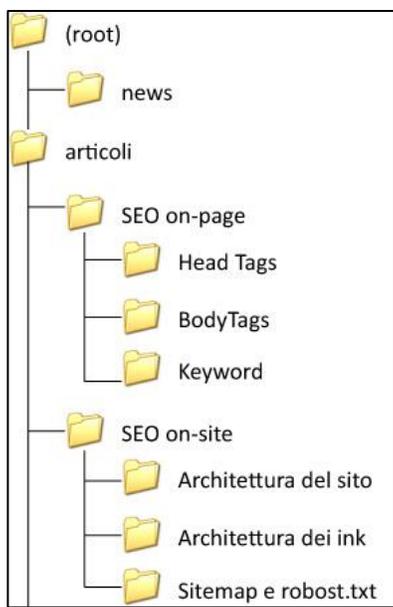


Figura 6.3 Schema ad albero di un sito

Partendo dalla *home* o “*root*” page, che in genere è la pagina più frequentata di un sito, e il punto di partenza della navigazione per molti utenti, bisogna pensare a come gli utenti passeranno da questa pagina a un’altra dal contenuto più specifico, e cosa essi si aspettano di trovare in queste pagine. Prima di tutto, dal punto di vista SEO, è buona norma inserire la parola chiave su cui si vuole puntare nel nome della cartella che contiene il file, perché offre agli spider informazioni migliori sulla struttura del sito e sulle sue pagine. Ovviamente, ogni pagina deve essere ottimizzata in tutte le sue parti, (title, description e body) per l’argomento che tratta, avere un link che rimanda alla home page, e link che rimandano alle altre sottopagine che trattano argomenti correlati. Dal punto di vista dell’utente invece, utili per facilitare la navigazione, e molto apprezzate, sono le cosiddette *breadcrumbs* (briciole di pane): una serie di link, in genere collocati all’inizio o alla fine della pagina, che comunicano costantemente all’utente dove si trova rispetto alla struttura del sito, permettendogli di navigare velocemente tra le sezioni precedenti o successive del sito, o di tornare direttamente alla home page. La figura 6.4 ne mostra un esempio.



Figura 6.4 Breadcrumbs

### 6.2.1 Usabilità e accessibilità del sito

L'usabilità di un sito, o *web usability*, è definita dall'ISO (International Organisation for Standardisation) come *l'efficacia, l'efficienza e la soddisfazione* con le quali determinati utenti raggiungono determinati obiettivi in determinati contesti<sup>32</sup>. L'obiettivo, è quindi quello di rendere intuitive per l'utente le azioni che è necessario svolgere sulle pagine, per raggiungere un certo risultato. Spesso, l'usabilità di un sito web e l'ottimizzazione per i motori di ricerca sono due aspetti complementari. Infatti, per essere efficace, è importante che un sito web si basi su standard riconosciuti, e una buona usabilità a sua volta, porta a vantaggi in termini di posizionamento nei risultati dei motori di ricerca.

Per facilitare la reperibilità delle informazioni e la visibilità delle pagine, nelle ancore dei link bisogna utilizzare testi semplici e contenenti le parole chiave principali, cercando di ridurre al minimo il numero di click necessari per raggiungere ciascuna pagina, specialmente quelle di maggiore importanza. Un altro aspetto importante per l'utente e per l'indicizzazione è l'ottimizzazione per il caricamento della pagina. Oggi nessuno ha pazienza, tantomeno nel web. Per cui, maggiore è il tempo di caricamento di una pagina, maggiori sono le probabilità che l'utente esca dal sito prima di averla visualizzata. Inoltre, un tempo di caricamento troppo lungo potrebbe portare gli spider ad abbandonare il sito, per passare ad analizzarne un altro. Bisogna quindi limitare l'uso di animazioni Flash e comprimere eventuali file multimediali per garantire un download più veloce.

Comunque, prima di decidere quale interfaccia utente utilizzare per il proprio sito, è meglio confrontare più alternative, eventualmente attraverso test in base ai profili demografici<sup>33</sup> degli utenti, e testarla per verificare che sia compatibile con i vari browser. La compatibilità tra i browser è un aspetto molto importante da tenere in considerazione, perché non tutti gli utenti utilizzano lo stesso browser, e soprattutto, non tutti i browser visualizzano le pagine allo stesso modo. Secondo una ricerca effettuata

---

<sup>32</sup>Marja-Riitta Koivunen and Matt May, *Exploring Usability Enhancements in W3C Process* [<http://www.w3.org/2002/09/usabilityws.html>]

<sup>33</sup> Termine utilizzato nel marketing per descrivere età, sesso, stato civile, orientamento religioso, cultura, posizione geografica, occupazione e altre metriche di segmentazione.

dall'azienda NetApplication<sup>34</sup> sulle quote di mercato dei browser, a Giugno 2011 Internet Explorer risultava ancora il browser più utilizzato dagli utenti, con il 53,78%, seguito da Mozilla Firefox con il 21,67% e da Google Chrome con il 13,11%. La tabella 6.1 illustra la il grafico completo della distribuzione sull'utilizzo dei browser.

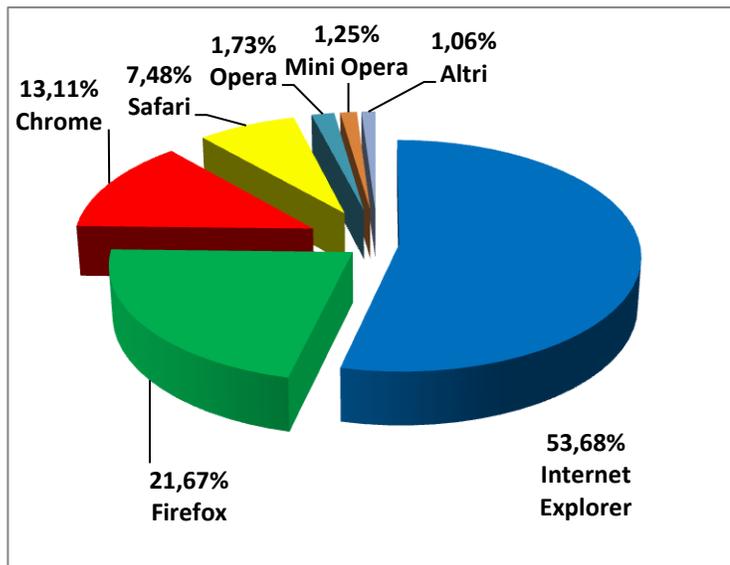


Tabella 6.1 Distribuzione dell'utilizzo dei browser

È importante inoltre tenere conto dei vari standard web e utilizzare validatori HTML per assicurare la migliore compatibilità. La maggior parte dei browser infatti, supportano varie versioni di HTML e dei protocolli web, ma pagine che non passano i test di validazione potrebbero venire indicizzate in modo strano dai motori di ricerca.

Per accessibilità, invece, si intende la possibilità da parte degli utenti disabili di percepire, capire, navigare e interagire con il web, e che possano inoltre contribuire al web<sup>35</sup>. Nonostante molti siti web non implementino tecniche di accessibilità, sarebbe opportuno, durante la fase di progettazione, pensare a questa categoria di persone che proprio nel web potrebbe trovare uno strumento vitale per comunicare, informarsi e ottenere servizi, portando a un sito web ulteriore traffico. Per questo, è stato introdotto

<sup>34</sup> Netmarketshare.com, *Browser Market Share* [<http://www.netmarketshare.com/browser-market-share.aspx?spider=1&qprid=0#>]

<sup>35</sup> Web Accessibility Initiative [<http://www.w3.org/WAI/intro/accessibility.php>]

uno standard consultabile sul sito della WAI (*Web Accessibility Initiative*) all'indirizzo <http://www.w3.org/WAI/>, dove si possono trovare le linee guida ufficiali per realizzare siti web accessibili (ATAG, *Authoring Tool Accessibility Guidelines*).

### **6.2.2 Architettura dei link**

L'architettura dei link interni è molto importante dal punto di vista della SEO, in particolare per siti con molte pagine o in continua crescita in termini di contenuti. Per creare una struttura ottimale per i link, come si è già detto nella sezione 6.2, spesso conviene suddividere i siti in sezioni e sottosezioni, in modo da avere una struttura gerarchica ad albero. Il problema più grande per l'indicizzazione di tutte le pagine di un sito, è però riuscire a dare risalto alle pagine più interne. Infatti, più sono i click per raggiungere una pagina, più questa pagina avrà difficoltà a essere indicizzata dagli spider, e di conseguenza, essere presente tra i risultati delle ricerche. A volte, inoltre, può capitare che uno spider non riesca a indicizzare a fondo un sito, a causa di qualche "trappola" che gli impedisce di seguire un link. Per poter evidenziare e risolvere problemi di questo tipo, basta navigare il sito con un browser testuale, come ad esempio *Lynx*, per capire come gli spider vedono una pagina web.

Creare una buona architettura dei link, aiuta quindi i motori di ricerca a trovare le pagine, aumentando le probabilità che esse vengano indicizzate. Inoltre, capire com'è distribuita la popolarità dei link interni, è importante perché permette di verificare il buon funzionamento di un sito in termini di SEO. Se la popolarità dei link interni cala mano a mano che ci sia allontana dalla home page, allora significa che c'è ancora del lavoro da fare. Un valido aiuto per fare individuare meglio ai motori di ricerca le pagine di un sito e facilitarne l'indicizzazione sono le Sitemap e i file robots.txt.

## 6.3 Sitemap XML e robots.txt

### Sitemap

Le Sitemap sono file in cui viene descritta in linguaggio XML la struttura delle pagine che compongono un sito web. Si rivelano molto utili ai fini della SEO, perché contengono le informazioni che aiutano il motore di ricerca a indicizzare correttamente i contenuti del sito, e aiutano gli utenti a raggiungere velocemente le informazioni di cui hanno bisogno. Inoltre, uno dei vantaggi nell'utilizzo delle Sitemap, deriva dalla possibilità di controllare la frequenza delle visite degli spider.

Come mostrato in figura 6.5, ogni link che compare nella sitemap può contenere quattro tag: <loc>, <lastmod>, <changefreq> e <priority>.

```
<?xml version='1.0' encoding='UTF-8'?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9
  http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd">
<url>
  <loc>http://miodominio.com/</loc>
  <lastmod>2011-01-01</lastmod>
  <changefreq>weekly</changefreq>
  <priority>0.5</priority>
</url>
</urlset>
```

Figura 6.5: Struttura di una sitemap XML

#### Il tag <loc>

E' l'unico tag obbligatorio e indica l'URL di destinazione del link.

#### Il tag <lastmod>

Indica la data e l'ora in cui l'URL è stato modificato. È un tag opzionale, ma molto potente, perché lo si può usare per segnalare agli spider che la pagina deve essere reindicizzata.

### **Il tag <changefreq>**

Indica la frequenza approssimativa di aggiornamento della pagina a cui si riferisce il link. È un tag opzionale e può assumere i seguenti valori: *always, hourly, daily, weekly, monthly, yearly, never*.

### **Il tag <priority>**

È il tag più importante. Può assumere valori compresi tra 0.0 e 1.0, il valore 0.5 è quello di default. Più alto è il suo valore, più aumenta la probabilità che l'URL venga visitato dagli spider. Tuttavia, l'assegnazione di una priorità 1.0 a tutti gli URL è controproducente e non porta a nessun vantaggio in termini di posizionamento.

Se non si è in grado di scrivere una sitemap, o se si ha un sito di grandi dimensioni, si può ricorrere a un generatore di sitemap XML. Uno strumento gratuito molto efficiente si trova a questo indirizzo: <http://www.xml-sitemaps.com/>. Una volta creata la sitemap, dovete assicurarvi che sia valida, altrimenti è inutile, e inviarla agli spider. Se si ha un account Google, con Google Webmaster Tools si può creare la sitemap e successivamente comunicarne ai motori di ricerca l'indirizzo.

### **Robots.txt**

L'utilizzo del file *robots.txt* è il metodo standard per comunicare agli spider cosa possono visitare e cosa no di un sito web. Questo file viene quindi usato anche per impedire ai motori di ricerca di indicizzare alcune parti di un sito, per esempio perché le pagine non sono ancora pronte, oppure perché non si vuole che certi contenuti compaiano nei risultati delle ricerche. Sebbene sia possibile includere in questo file anche le regole d'accesso valide solo per singole pagine, è comunque preferibile inserire il meta tag *robots* direttamente all'interno del codice HTML delle pagine stesse. Infatti, i vari spider non interpretano esattamente allo stesso modo le regole di accesso che leggono dal file *robots.txt*, poiché i motori di ricerca ancora non hanno trovato un accordo sullo standard. L'utilizzo del file non è obbligatorio, però si consiglia di utilizzarlo, visto che può rivelarsi molto utile nella gestione dei contenuti duplicati.

Per creare un *robots.txt* basta un qualsiasi editor di testi per inserirvi le regole di accesso. Una volta creato il file, bisogna assegnargli permessi in sola lettura, in modo che possa

avere un accesso libero, ma non sia possibile modificarlo. Deve risiedere nella cartella principale del sito a cui si riferisce, e infine, poiché deve essere sintatticamente e funzionalmente valido, bisogna convalidarlo. Sul web esistono molti validatori gratuiti, ma anche su Google Webmaster Tools se ne può trovare uno.

Una considerazione importante da fare, è che ogni motore di ricerca utilizza propri spider, e tali spider non sono tutti uguali. Alcuni sono progettati per visitare le pagine web, altri solo le immagini, i feed RSS, i video e così via. Per cui, a seconda dei contenuti del sito e del target geografico di riferimento, conviene decidere quali spider gestire nel file *robots.txt*, in modo da consentire agli spider di analizzare i contenuti che si vogliono indicizzare, favorendo notevolmente l'ottimizzazione del sito. Cercando su Google `file type:text robots.txt`, si possono trovare molti esempi di configurazione di questo file. Ad esempio, si può specificare di bloccare l'indicizzazione delle immagini, di documenti o impedire l'ingresso a determinati spider. Nella figura 6.6 è mostrato un esempio che definisce le regole per tre diversi spider.

```
1 User-agent: *
2 Disallow: /
3 Allow: /articles/
4 Allow: /news/
5 Allow: /blog/
6
7 User-agent: msnbot
8 Disallow:
9 User-agent: googlebot
11 Disallow: cgi-bin/
```

Figura 6.6: Struttura di un file robots.txt

Le prime cinque righe definiscono le regole di default valide per tutti gli spider. Il carattere speciale asterico \* infatti, ha un significato simile a quello utilizzato nelle espressioni regolari, perché indica zero o più occorrenze di un carattere in una sequenza, e quindi indica le direttive per *tutti* gli spider che visiteranno il sito. Le sezioni successive invece, riguardano gli spider di Google e Bing: *googlebot* e *msnbot*. La

direttiva `Disallow` specifica un percorso in cui gli spider non possono entrare, mentre la direttiva `Allow` specifica un percorso per cui gli spider sono autorizzati. Quando ad esempio *Slurp*, lo spider di Yahoo! entra in questo sito, seguirà solo il primo gruppo di regole, poiché il file non contiene nessuna sezione che lo riguarda. Per avere un quadro degli spider utilizzati dai motori di ricerca, nella figura 6.7, sono illustrati quelli principali attualmente in circolazione sul web.

Google	
Spider	Descrizione
Googlebot	Indicizza le pagine web (il più importante in assoluto)
Googlebot-Mobile	Indicizza le pagine destinate alla visualizzazione sui dispositivi mobili
Googlebot-Image	Indicizza le immagini
Mediapartners-Google	Visita le pagine che utilizzano il servizio AdSense
AdsBot-Google	Visita le pagine associate alle parole chiave che utilizzano il servizio AdWords
Yahoo!:	
Spider	Descrizione
Slurp	Indicizza le pagine web
Yahoo-MMAudVid	Indicizza i contenuti video
Yahoo-MMCrawler	Indicizza le immagini
Bing:	
Spider	Descrizione
MSNBot	Indicizza le pagine web
MSNBot-Media	Indicizza i contenuti multimediali
MSNBot-News	Indicizza i fed RSS

Figura 6.7 I principali spider dei motori di ricerca

## 7. Ottimizzazione off-site

In questa sezione, verranno illustrati i fattori esterni ad un sito web che influenzano il posizionamento. Su di essi non si può esercitare un controllo diretto, poiché riguardano ad esempio il numero di link in entrata, i modelli di comportamento degli utenti, le prestazioni del sito o l'età del dominio.

### 7.1 Popolarità

Tra i fattori esterni attualmente più importanti che influenzano il posizionamento sui motori di ricerca c'è la popolarità, una misura che indica l'affidabilità dei contenuti di un determinato sito web<sup>36</sup>. Più un sito è ritenuto affidabile dai motori di ricerca, e più questi lo posizionano in alto tra le SERP. All'inizio, i motori di ricerca erano in grado di comprendere la popolarità di un sito calcolando i link che questo riceveva dagli altri siti. Con questo semplice concetto, tutti approfittarono dei vantaggi che la cosiddetta *link popularity* portava a un sito in termini di visibilità. Questo sistema però, favorì anche l'attività degli spammer, che iniziarono a infestare le SERP con milioni di pagine spam<sup>37</sup>, causando il problema relativo alla cattiva qualità dei contenuti nei risultati delle ricerche. Nel 1998 però, Larry Page e Sergey Brin svilupparono un nuovo concetto di link popularity, che fu testato su un nuovo motore di ricerca: Google. Il nuovo concetto di link popularity, chiamato PageRank<sup>38</sup>, portò Google alla ribalta come migliore motore di ricerca, generando un enorme distacco qualitativo tra Google e gli altri motori. La link popularity sviluppata da Google non prende in considerazione solo il numero di link, ma assegna a un documento presente nel proprio database un punteggio numerico attraverso gli stessi link, valutandone la qualità e diversificandone l'influenza. Maggiore è il valore del documento, maggiore quel link è considerato importante e influente ai fini del posizionamento. Automaticamente, maggiore è il numero di link in ingresso con un

---

<sup>36</sup> Wikipedia, voce *Link popularity* [[http://it.wikipedia.org/wiki/Link\\_popularity](http://it.wikipedia.org/wiki/Link_popularity)]

<sup>37</sup> Pagine web più o meno fittizie, con contenuti non rilevanti per l'utente e finalizzate a ottenere accessi

<sup>38</sup> Page L., Brin S. *Method for node ranking in a linked database* [<http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetacgi%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=6,285,999.PN.&OS=PN/6,285,999&RS=PN/6,285,999>]

valore alto verso un determinato sito, maggiore è la popolarità che questo sito acquisisce agli occhi dei motori di ricerca. Quindi, è importante riuscire a far guadagnare a un sito in modo spontaneo link su altri siti, che siano a loro volta popolari, in modo da far diventare il sito autorevole agli occhi dei motori di ricerca e quindi degno delle prime posizioni tra i risultati delle ricerche.

### 7.1.1 Link esterni

I link esterni, chiamati anche *backlink*, *link in ingresso* o *referral link*, sono appunto link che permettono di raggiungere un sito, o una pagina di un sito, da altri siti. Come si è detto, dal punto di vista della SEO il numero di link esterni sono un'indicazione della popolarità o importanza di un sito web o di una pagina. E' importante perciò considerare quali sono questi siti che ci linkano, specialmente quando la popolarità aumenta, e stabilire il giusto equilibrio tra quantità e qualità dei backlink. Per sapere chi ha un link verso un determinato sito, basta andare su Google e utilizzare il comando `link:` seguito dal nome del sito che si vuole controllare, oppure, si può andare su Google Webmaster Tools, dove è stata introdotta una funzione analoga. Stessa operazione si può fare anche con Yahoo! Site Explorer, con il quale si possono esportare in un file con formato TSV i primi 1.000 backlink verso il sito in questione. Molti link in ingresso sono sicuramente importanti, ma altrettanto importante è che questi siano di qualità. La cosa migliore infatti sarebbe avere link provenienti da siti molto popolari, perché sono sinonimo della fiducia accordata al sito che riceve il link, e i motori di ricerca basano sempre più i propri algoritmi di posizionamento proprio sulla fiducia di cui gode un sito. I benefici dei link in ingresso sono duplici per un sito web, perché riceve traffico dagli altri siti che lo linkano e dai motori di ricerca.

Per ottimizzare al meglio i backlink, è importante che contengano nelle ancore dei testi rilevanti, devono perciò essere pertinenti e contenere, come sempre, la parole chiave principale. Inoltre, sarebbe opportuno posizionarli in cima alla pagina HTML, in modo da aumentarne la prominenza. Bisogna ricordare poi, che l'efficacia di un backlink non è immediata, per cui, se non si ottengono subito i risultati sperati non bisogna preoccuparsi.

### 7.1.2 PageRank

Il PageRank (chiamato comunemente PR) è l'algoritmo sviluppato da Google nel 1998, e che costituisce uno dei più importanti fattori per stabilire il posizionamento di un sito web nelle SERP di questo motore di ricerca. Il nome deriva dall'unione delle parole Page e Rank. Page è il nome di uno dei due inventori, e allo stesso tempo significa *pagina* in inglese, e Rank, che sta per *classifica*. Questo algoritmo assegna a ogni documento (pagina) presente nel database di un motore un'importanza, assegnando a ogni link presente nella pagina un peso numerico, con lo scopo di quantificarne la sua importanza tra tutti i documenti presenti nella collezione<sup>39</sup>. Questo valore assegnato, è calcolato in base ai valori delle pagine che attraverso i link portano ad esso. Più il documento che ci linka ha un valore alto, più sarà importante per una determinata pagina questo link. La formula completa per calcolare il PageRank è la seguente:

$$PR(A) = (1 - d) + d \left( \sum_{k=1}^n \frac{PR[P_k]}{C[P_k]} \right)$$

Dove:

- $PR(A)$  rappresenta il PageRank della pagina  $A$ ;
- $\frac{PR[P_k]}{C[P_k]}$  rappresenta il rapporto tra il PageRank della pagina  $P_k$  che ha un backlink verso la pagina  $A$ , e il numero di link in uscita  $C[P_k]$  che contiene la pagina  $[P_k]$ ;
- $d$  è il *damping factor* cioè un valore compreso tra 0 e 1 che Google assegna ad un sito appena viene pubblicato.

Quindi la somma di tutti i valori  $\frac{PR[P_k]}{C[P_k]}$  moltiplicati per  $d$  e sommati a  $(1 - d)$  danno come risultato il PageRank della pagina  $A$ .

---

<sup>39</sup> Wikipedia, voce *Page Rank* [<http://it.wikipedia.org/wiki/PageRank>]

Per vedere il PageRank di una pagina, possiamo affidarci alla *toolbar* di Google (scaricabile all'indirizzo <http://toolbar.google.it>), oppure ai molti strumenti gratuiti disponibili sul web. Bisogna dire però, che questo algoritmo per quanto efficace, non è in grado di eliminare totalmente le pagine spam. Quindi, per migliorare sensibilmente la qualità dei risultati nelle ricerche, Google ha implementato un altro algoritmo, chiamato TrustRank.

### 7.1.3 TrustRank

L'algoritmo di TrustRank consiste nell'identificare una serie di pagine (dette pagine seme), delle quali non è ancora certo se presentino o meno dello spam<sup>40</sup>. Le pagine seme vengono individuate basandosi su un semplice concetto: un sito di alta qualità difficilmente linkerà un sito di bassa qualità. Un controllo manuale assegna loro un punteggio, 0 se la pagina presenta dello spam, 1 se è considerata buona. In seguito, a tutti i documenti presenti nel database del motore, l'algoritmo assegna un valore di *trust* (fiducia) a seconda della vicinanza (mediante i link) di questi documenti con le pagine seme che risultano prime di spam. Le caratteristiche su cui si basa il TrustRank sono sostanzialmente due, e sono state chiamate *Trust dampening* e *Trust splitting*.

#### Trust dampening

Il Trust dampening (smorzamento della fiducia) si basa su questo concetto: più un documento si allontana dalla pagina seme, più aumenta la probabilità di trovare un documento che contiene spam. Nella figura 7.1 è illustrato lo schema che lo rappresenta.

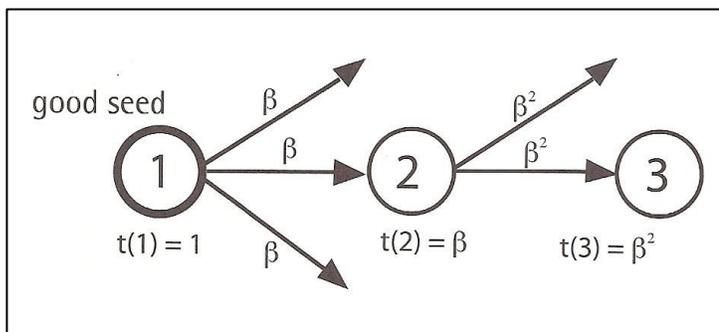


Figura 7.1 Schema del Trust dampening

<sup>40</sup> Gyöngyi Z., Garcia-Molina H., Pedersen J., *Combating Web Spam with TrustRank* [<http://www.vldb.org/conf/2004/RS15P3.PDF>]

Nella figura, la pagina 1 appartiene all'insieme di pagine seme buone, e contiene un link che punta alla pagina 2, alla quale passa un valore di fiducia  $\beta$  minore di 1. Alla pagina 3, che invece è raggiungibile direttamente dalla pagina 2, viene trasmesso un valore di fiducia uguale a  $\beta*\beta$  e così via. Nel caso in cui le pagine ricevano fiducia da link multipli, può essere assegnato ad esse il valore maggiore trasmesso da una singola pagina oppure una media di tutti i valori.

### Trust splitting

Il Trust splitting (suddivisione della fiducia) invece, è un valore che si basa sulla cura e sull'attenzione dedicate all'inserimento nelle pagine di un determinato sito web, dei cosiddetti *outbound link*, o link in uscita, ovvero link di altri siti. Questo valore, viene considerato inversamente proporzionale al numero totale dei link in uscita presenti in un sito. Se una pagina buona contiene pochi link, questi molto probabilmente punteranno a pagine a loro volta buone, mentre se la stessa pagina contiene centinaia di link è più probabile che alcuni di essi puntino a pagine che contengono spam. La figura 7.2 illustra lo schema relativo al funzionamento del Trust splitting.

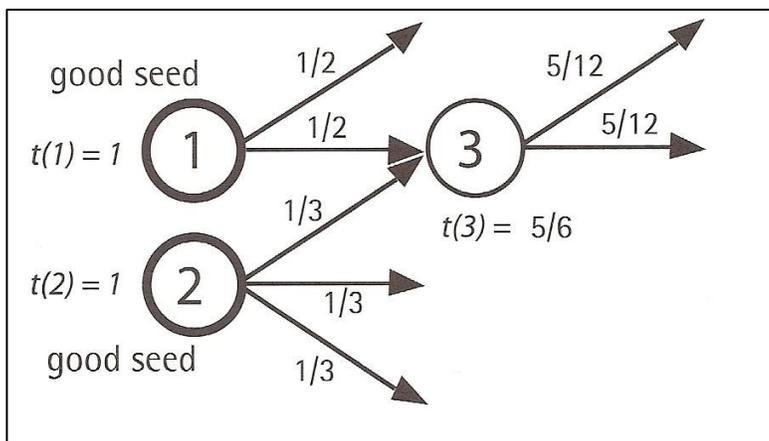


Figura 7.2 Schema del Trust splitting

Se una pagina buona ha un valore di fiducia  $t$  e contiene  $\omega$  link ad altre pagine, ad ognuna di queste sarà trasmesso un valore di fiducia uguale a  $t/\omega$ . In questo caso, i punteggi che una pagina riceve da link multipli provenienti da pagina buone vengono sommati.

Nella figura, la pagina 1, appartenente all'insieme di pagine seme buone, contiene due link in uscita, così assegna a ciascuna delle pagine a cui punta un valore pari a 0,5 (la metà della sua fiducia). Anche la pagina 2 appartiene all'insieme di pagine seme buone S, ma contiene tre link in uscita, quindi trasmette a ciascuna delle pagine a cui punta un valore pari a 0,333 (un terzo della sua fiducia). La pagina 3, riceverà quindi una fiducia totale pari a  $0,5+0,333=0,833$ .

Il Trust dampening e il Trust splitting possono anche essere combinati. In questo caso, sempre riferendosi alla fig.7.2, la pagina 3 riceverà un punteggio di  $\beta*(0,5+0,333)$ . Così come per il PageRank, anche per il TrustRank, e in particolare per l'aspetto che riguarda il Trust splitting, la quantità di link esterni non abbassa necessariamente il valore di una determinata pagina. Al massimo, si passerà meno valore ai documenti che vengono linkati, e inoltre, inserire link a siti di buona qualità rappresenta un contributo per le ricerche degli utenti, diventando di conseguenza un fattore molto positivo per un sito. Per calcolare e quantificare il TrustRank, non esiste nessuno strumento. Tuttavia, trovare siti con alti valori di fiducia è molto semplice. Basta cercare su Google le chiavi di ricerca che ci interessano, e i primi venti siti che il motore di ricerca restituirà tra le sue ricerche avranno sicuramente un buon TrustRank.

## 7.2 Come aumentare la popolarità

Uno dei fattori che incide di più sulla popolarità di un sito, come si è detto, è la qualità dei link che provengono da altri siti web. Per aumentare la link popularity, bisogna capire anche come i motori di ricerca valutano questi link. Come si è detto infatti, non tutti i backlink sono uguali, alcuni portano vantaggi e altri invece possono essere penalizzanti. I motori di ricerca, in genere, cercano di dare importanza solo ai link che veramente creano un vantaggio all'utente, cercando di capire anche se un link è spontaneo, e quindi rappresenta una vera risorsa per l'utente, oppure se è il frutto di una compravendita, di uno scambio o comunque di accordi con il solo fine di accrescere la popolarità di un sito. Oltre al *fattore qualità*, bisogna considerare anche il *fattore tempo*. È importante che un sito riceva backlink costantemente e in modo progressivo. Con una crescita naturale dei backlink infatti, i motori di ricerca tenderanno a interpretare un

certo sito come importante per gli utenti, e quindi degno di fiducia. Comunque, per facilitare l'aumento di popolarità, sono state sviluppate diverse tecniche, come lo scambio di link e l'utilizzo delle directory. Fino a qualche tempo fa, queste tecniche erano molto valide per migliorare il posizionamento. I motori di ricerca però si sono resi conto del loro utilizzo eccessivo, e adesso il valore dello scambio di link e delle directory è praticamente nullo. Attualmente le tecniche più valide per aumentare la popolarità, sono i social bookmarking e i feed RSS.

### **7.2.1 Social bookmarking**

I *social bookmarking* sono siti che offrono la possibilità di pubblicare sul web gli URL preferiti da un utente, mettendoli a disposizione degli altri utenti della comunità, come per esempio *Delicious*. Se, partecipando a una web community, un utente individua contenuti o pagine particolarmente interessanti, tramite i social bookmarking si possono condividere subito con gli altri utenti. I vari social bookmarking offrono funzioni diverse tra loro, ma le più significative permettono di:

- rendere il documento visibile per tutti quegli utenti che sono interessati a quell'argomento specifico;
- inserire il documento nel database dei motori di ricerca, rendendo la pagina visibile nelle SERP, e portando così più traffico al sito;
- offrire ai propri utenti la possibilità di votare il *bookmark*, ovvero quel suggerimento, quando è ritenuto particolarmente interessante, aumentando così la popolarità del sito.

Questi siti sono molto utili anche per allargare il bacino di utenza a cui ci si rivolge. Offrendo contenuti di qualità targhettizzati, infatti, aumenta la probabilità di conversione da parte dell'utente, portando di conseguenza al sito un traffico costante e progressivo, incidendo sulla popolarità anche agli occhi dei motori di ricerca.

## 7.2.2 Feed RSS

RSS è l'acronimo di *Really Simple Syndication*. La syndication è lo standard utilizzato per la distribuzione automatica di contenuti aggiornati sul web<sup>41</sup>. Chi pubblica contenuti li rende disponibili in vari formati, e la syndication mette a disposizione un mezzo per far raggiungere questi contenuti a un ampio numero di utenti. Se un sito offre questo servizio, compare una piccola icona, come mostrato in figura 6.9.



Figura 6.9 Icona Feed RSS

Per abbonarsi a un feed RSS è semplicissimo: cliccando sull'icona, il feed verrà automaticamente aggiunto all'aggregatore impostato come predefinito e ogni volta che un sito pubblica contenuti nuovi, questi saranno pubblicati nell'aggregatore.

Un feed è un semplice file XML e per poterlo interpretare, e rendere quindi fruibile i contenuti, ci sono diverse modalità. Le più diffuse sono due: attraverso appositi software, da installare sul proprio computer, oppure attraverso aggregatori web, chiamati *feed reader*. Questi ultimi si rivelano molto utili, in quanto sono in grado di visualizzare i contenuti di un feed provenienti da diversi siti. In questo modo, l'utente può accedere a tutte le notizie direttamente da un unico sito web (aggregatore), senza dover visitare ogni singolo sito da cui proviene la notizia. In questo modo, si aumenta ulteriormente il bacino di utenza di una determinata notizia e la possibilità che altri siti redistribuiranno a loro volta i contenuti del sito a cui si sono abbonati.

<sup>41</sup> Wikipedia, voce RSS [<http://en.wikipedia.org/wiki/RSS>]

Altri due fattori che non sono strettamente collegati alle tecniche SEO, ma che meritano costituiscono un importante fattore per aumentare la popolarità di un sito, sono i contenuti e i social network.

#### **7.2.4 I contenuti**

“Content is the king”. Questo è quello che spesso si afferma tra i professionisti SEO. Infatti, uno dei modi migliori per aumentare la popolarità di un sito e ricevere molti backlink spontanei, è pubblicare contenuti di qualità. Contenuti unici e originali, aggiornati frequentemente e fruibili in più formati (HTML, PDF, RSS, video ecc.) costituiscono la chiave principale per il successo di un sito. È fondamentale inoltre creare contenuti ritenuti utili e che siano apprezzati dagli utenti a cui ci si rivolge. Dal punto di vista della SEO, i contenuti sono un aspetto molto importante, perché è lì che si compie l'utilizzo strategico delle parole chiave. In un sito, si possono distinguere due categorie principali di contenuti: contenuti a breve termine e contenuti a lungo termine.

##### **Contenuti a breve termine**

I contenuti a breve termine sono quelli che inseguono le mode del momento e sono importanti per dare al sito un'immagine di freschezza e tempestività. Per ottenere buoni risultati in un breve periodo, però, bisognerebbe riuscire ad anticipare queste tendenze. Lo strumento di Google Trends, analizzato nella sezione 4.2, fornisce un valido supporto per capire come sul web cambiano le mode. Infatti, sapere quali sono gli schemi d'uso delle parole chiave, aiuta anche a farsi un'idea di quello che possono essere le tendenze del futuro.

##### **Contenuti a lungo termine**

I contenuti a lungo termine sono quelli destinati a durare nel tempo. Sono quelle pagine che gli utenti aggiungono ai propri bookmark e che vengono visualizzate molte volte. Questo tipo di contenuti può riguardare qualsiasi aspetto, e la loro rilevanza non diminuisce nel tempo. Gli utenti infatti, quando cercano informazioni di riferimento, continuano la loro ricerca fino a che non trovano esattamente quello che stanno

cercando. Per cui, questo è il caso in cui avere siti dedicati e con contenuti di qualità può fare la differenza e aumentare il traffico e la popolarità di un sito.

### **7.2.3 Social network**

I *social network* hanno introdotto nel mondo del web numerosi cambiamenti. Essi, infatti, incidono nel calcolo della link popularity perché, mettendo a disposizione ulteriori canali, riescono a veicolare molto traffico sui siti web. La loro utilità è così importante, che è stato introdotto un settore specifico chiamato *SMO*, Social Media Optimisation, in grado di aiutare un sito a costruire un'immagine di fiducia, influenza e autorevolezza. Dal punto di vista della SEO, quindi, raggiungere una notorietà sui social network significa sviluppare in modo naturale link in ingresso a un determinato sito, aumentandone di conseguenza la popolarità. Uno dei maggiori problemi SEO riguardo alla condivisione di contenuti sui social network, è l'attributo `rel="nofollow"`, utilizzato per comunicare agli spider di non indicizzare un determinato link. Quasi tutti i social network, infatti, seguono il comando `nofollow`. Quindi, i link che vi vengono pubblicati, non influiscono per il calcolo del PageRank perché non sono presi in considerazione dai motori di ricerca. Non bisogna comunque focalizzarsi su questo aspetto, perché un link pubblicato può portare comunque traffico e conversioni.

## 8. Caso di studio: [www.toscanaper.me](http://www.toscanaper.me)

*Toscanaper.me* è un blog ideato per posizionare Mediaus, l'azienda presso la quale ho svolto il tirocino, all'interno del settore turistico in Toscana, con un progetto editoriale di buon livello. Scopo del blog, è di fornire informazioni a tutti coloro che sono interessati a visitare la Toscana e a scoprire cosa questa regione offre durante tutto il corso dell'anno. Eventi, itinerari, attività, cultura, territorio ed enogastronomia sono le sei categorie che sono state scelte come punto di riferimento per le nostre proposte. Il target di riferimento è ampio, ma con una preferenza verso argomenti di interesse prevalentemente giovanile. La scelta del blog è giustificata da un approccio tipicamente Web 2.0: commenti sul sito e condivisione dei post sui principali social network, vogliono essere il punto di partenza per rendere *toscanaper.me* un punto di riferimento e di interazione, attraverso lo scambio di idee e consigli, per aiutare i visitatori con informazioni utili e i potenziali clienti del settore turistico dell'azienda, a farsi conoscere e promuovere prodotti e servizi. L'obiettivo finale è, infatti, quello di diventare un *brand* per gli amanti della Toscana e per i toscani stessi. Per farlo, *toscanaper.me* ha dato importanza, in questa fase iniziale, soprattutto alla generazione di traffico dai motori di ricerca, e meno all'aspetto della monetizzazione in tempi brevi con le campagne pubblicitarie di Google AdSense. Questo riflette difatti l'obiettivo finale: ottenere una buona visibilità dai motori di ricerca e una buona posizione tra i risultati delle ricerche nell'ambito del web marketing turistico. Per quanto riguarda le campagne AdSense, queste sono presenti, ma non rappresentando l'obiettivo primario, sono state posizionate all'interno del sito in spazi dedicati poco invasivi per l'utente. Da poco, dato l'aumento dei volumi di traffico, *toscanaper.me* è stata lanciata sul social network Facebook per aumentare la popolarità e favorire la fidelizzazione dell'utente.

## 8.1 CMS Wordpress

Il sito *toscanaper.me* è stato realizzato mediante il CMS<sup>42</sup> Wordpress, versione 3.2.1. La ragione principale di questa scelta è che Wordpress rappresenta il migliore CMS per la gestione di blog e progetti editoriali di medio e alto livello. Inoltre è gratuito, il sistema viene aggiornato frequentemente, e vanta una enorme comunità di sviluppatori che producono plug-in per migliorarne le funzionalità. Questo ultimo aspetto si è rivelato particolarmente utile dal punto di vista SEO, infatti, con Wordpress non è possibile compiere alcune funzioni per migliorare l'ottimizzazione del sito e delle singole pagine, come la gestione dei meta tag description, keyword e robots, quindi, è stato necessario implementare alcuni plug-in come: *Platinum SEO Pack*, *Google XML Sitemap*, *Hobo WP Googlebot Tracker* e *W3 Total Cache*.

### Platinum SEO Pack

Questo plug-in offre una soluzione completa per l'ottimizzazione on-site di un blog Wordpress. Come mostrato in figura 9.1, tra le funzionalità più importanti troviamo la possibilità di personalizzare il tag `title` e i meta tag `description` e `keyword` in ogni pagina, la possibilità di gestire la riscrittura delle URL, di impostare l'URL canonico e di scegliere quali eventuali contenuti non far seguire dagli spider mediante la gestione dei metadati `noindex` e `nofollow`.

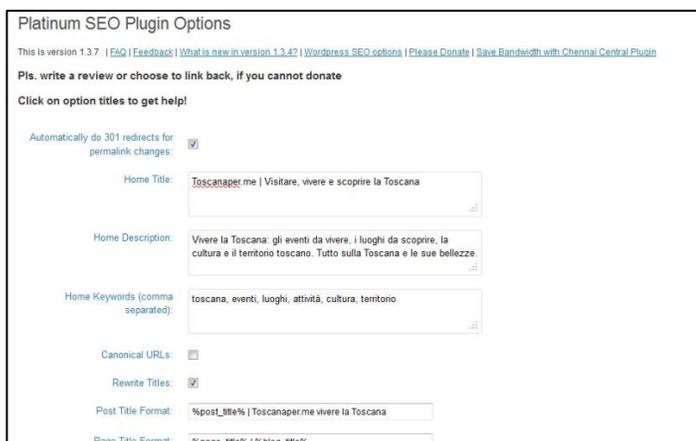
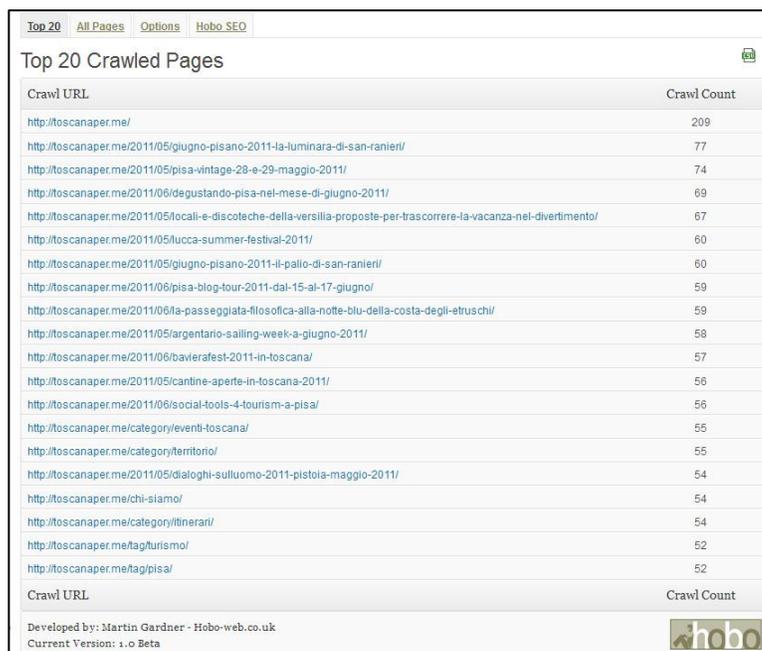


Figura 9.1 Platinum SEO Pack plug-in per Wordpress

<sup>42</sup> Content Management System (Sistema di Gestione dei Contenuti) è un software installato su un server web studiato per permettere la gestione dei contenuti di un sito web, senza aver bisogno di conoscere le tecniche di programmazione.

## Hobo WP Googlebot Tracker

Questo plug-in è utilizzato per monitorare quali pagine vengono visitate dagli spider di Google. Inoltre, fornisce informazioni circa l'ora e il giorno in cui è stata effettuata la scansione, e la tipologia di spider Googlebot. La figura 9.2 mostra le 20 pagine di *toscanaper.me* più visitate dagli spider.



The screenshot shows the 'Top 20 Crawled Pages' section of the Hobo WP Googlebot Tracker plugin. It features a table with two columns: 'Crawl URL' and 'Crawl Count'. The top 20 entries are listed below, with the highest crawl count being 209 for the root URL.

Crawl URL	Crawl Count
http://toscanaper.me/	209
http://toscanaper.me/2011/05/giugno-pisano-2011-1a-luminara-di-san-ranieri/	77
http://toscanaper.me/2011/05/pisa-vintage-28-e-29-maggio-2011/	74
http://toscanaper.me/2011/06/degustando-pisa-nel-mese-di-giugno-2011/	69
http://toscanaper.me/2011/05/focall-e-discoteche-della-versilia-proposte-per-trascorrere-la-vacanza-nel-divertimento/	67
http://toscanaper.me/2011/05/lucca-summer-festival-2011/	60
http://toscanaper.me/2011/05/giugno-pisano-2011-ii-palio-di-san-ranieri/	60
http://toscanaper.me/2011/06/pisa-blog-tour-2011-dal-15-al-17-giugno/	59
http://toscanaper.me/2011/06/la-passeggiata-filosofica-alla-notte-blu-della-costa-degli-etruschi/	59
http://toscanaper.me/2011/05/argentario-sailing-week-a-giugno-2011/	58
http://toscanaper.me/2011/06/baviefest-2011-in-toscana/	57
http://toscanaper.me/2011/05/cantine-aperte-in-toscana-2011/	56
http://toscanaper.me/2011/06/social-tools-4-tourism-a-pisa/	56
http://toscanaper.me/category/eventi-toscana/	55
http://toscanaper.me/category/territorio/	55
http://toscanaper.me/2011/05/dialoghi-sull'uomo-2011-pistola-maggio-2011/	54
http://toscanaper.me/chi-siamo/	54
http://toscanaper.me/category/itinerari/	54
http://toscanaper.me/tag/turismo/	52
http://toscanaper.me/tag/pisa/	52

Developed by: Martin Gardner - Hobo-web.co.uk  
Current Version: 1.0 Beta

Figura 9.2 Hobo WP Googlebot Tracker plug-in

## W3 Total Cache

Questo è il più completo e complesso plug-in per Wordpress, ed è utilizzato per migliorare la *user experience* dell'utente, migliorando le prestazioni del server e ottimizzando, di conseguenza, il tempo di caricamento delle pagine e dei suoi elementi. Il vantaggio più importante dal punto di vista SEO che deriva dall'utilizzo di questo plug-in, è il miglioramento del ranking delle pagine su Google, e il conseguente aumento della probabilità che gli utenti trascorrono più tempo sul sito.

Infine, per migliorare l'ottimizzazione del titolo di ogni post, sono stati utilizzati i *permalink* (abbreviazione di *permanent link*), un tipo di URL associata ad una determinata informazione, implementata in modo da rimanere invariata per sempre o per

molto tempo<sup>43</sup>. I permalink vengono usati spesso nell'ambito dei blog per indicare il link a un determinato post, e sono molto utili per due fattori: cambiando la struttura del permalink, si rendono più comprensibili le URL agli utenti, passando da una URL dinamica (con parametri) a una URL statica. Allo stesso tempo, queste URL sono più comprensibili anche per i motori di ricerca, facilitando così l'indicizzazione dei singoli post.

La struttura di un permalink può essere di diverso tipo, dipende dalle preferenze o dalle esigenze del blog, ma *toscanaper.me* ha utilizzato quella composta dall'anno e il mese di pubblicazione seguito dal titolo del post: `/%year%/%monthnum%/%postname%/`. In questo modo, i motori di ricerca sono in grado di decodificare le informazioni relative alla data nella URL, e di visualizzarla poi accanto al risultato nella SERP, mentre l'utente è in grado di verificare se una determinata pagina è recente o meno, aumentando le possibilità che visiti il sito.

## 8.2 Analisi delle parole chiave

Le categorie del menù sono state il punto di partenza per la ricerca e l'analisi delle parole chiave. Intorno a queste, sono state poi scelte delle keyword generiche da utilizzare come tag. L'utilizzo dei tag e del *tag cloud*<sup>44</sup>, infatti, può essere utile sia all'utente per individuare le parole chiave più importanti all'interno del sito, e facilitare la navigazione basata sulle keyword, sia al posizionamento delle pagine dal punto di vista SEO, perché indica la frequenza di utilizzo e il peso di una certa keyword nel sito. Poiché una lista di tag troppo lunga potrebbe causare il keyword stuffing (v. sezione 5.3), danneggiando il posizionamento, per *toscanaper.me* sono state scelte circa trenta parole chiave mirate, che riguardano le principali città della Toscana e alcune parole che classificano in modo più specifico gli argomenti trattati nei post delle varie categorie, orientando meglio gli utenti. Queste sono: *Arezzo, arte, benessere, campagna, Carrara, Chianti, estate, Firenze, Garfagnana, Grosseto, Lucca, luoghi, mare, Maremma, Massa,*

---

<sup>43</sup> Wikipedia, voce *Permalink* [<http://it.wikipedia.org/wiki/Permalink>]

<sup>44</sup> Tradotto significa "nuvola di tag". È una rappresentazione visiva delle parole-chiave usate in un sito web mediante una lista, generalmente in ordine alfabetico. La caratteristica principale è che maggiore è l'importanza di una parola nel sito o in una pagina specifica, maggiore è il font con cui essa viene rappresentata.

*montagna, musica, Pisa, Pistoia, Prato, ristoranti, Siena, sport, turismo, vacanza, Versilia, weekend.*

Successivamente, mediante Google AdWords Keyword Tool, per ogni categoria è stato stilato un elenco di parole chiave su cui iniziare a lavorare, visibile nell'appendice A, partendo da quelle più generiche e andando via via a restringere il campo con keyword più specifiche, puntando prevalentemente sulle parole chiave *long tail*. Per la scelta, sono stati valutati i tre parametri principali dello strumento Google: livello di concorrenza, volume di ricerca mensile globale e volume di ricerca mensile locale. Inoltre, per avere risultati più precisi, è stata delimitata la ricerca al tipo di corrispondenza esatta. Per quanto riguarda la categoria "eventi", invece, la scelta delle keyword è stata abbastanza naturale, perché queste corrispondono al nome dell'evento, al luogo e alla data (giorno o intervallo di giorni) in cui l'evento si svolge.

Infine, per ogni keyword di ogni singolo argomento, sono state cercate parole correlate e derivate della keyword stessa, per evitare lo sgradevole effetto della ripetizione meccanica delle parole chiave nel testo e, allo stesso tempo, aumentare la pertinenza della pagina agli occhi dei motori di ricerca.

## 8.3 Ottimizzazione

Per *toscanaper.me*, l'attenzione è stata incentrata soprattutto sull'ottimizzazione on-page, in modo da facilitare l'indicizzazione degli articoli pubblicati e aumentare rapidamente la visibilità del sito. Il lavoro da me svolto, è stato quello di scrivere articoli sulla Toscana che fossero interessanti dal punto di vista dei contenuti e, allo stesso tempo, che fossero ben strutturati dal punto di vista SEO. Un valido aiuto è stato fornito dal Platinum SEO Pack, che ha permesso di personalizzare i tag `title` e i `meta tag` di ogni articolo perché, come si è visto, sono estremamente utili per i motori di ricerca. Inoltre, per facilitare il posizionamento, prima della pubblicazione ogni articolo è stato sottoposto a revisione, in modo da migliorare la prominenza, la prossimità e la densità delle keyword nel testo, facendo comunque attenzione a non incorrere nel keyword stuffing. Anche le immagini sono state ottimizzate: le dimensioni sono state alleggerite per non rallentare il caricamento della pagina, ma senza penalizzare la risoluzione, il nome dei file, i titoli delle immagini e l'attributo `alt` sono stati scritti in modo chiaro e descrittivo inserendo nel testo la keyword principale, in modo da aiutare il posizionamento tra i risultati delle ricerche relative alle immagini.

Per quanto riguarda l'ottimizzazione on-site invece, si è data molta importanza all'organizzazione interna dei link, per aiutare i motori di ricerca a trovare le varie pagine del sito e indicizzarle. Nella home page e nella pagina principale delle varie categorie, si trovano i link agli ultimi cinque articoli pubblicati, accompagnati da una breve anteprima dei contenuti dell'articolo. In ogni pagina inoltre, nella prima colonna di destra troviamo i link agli articoli più recenti, insieme ai link della *tag cloud* e quelli delle categorie. Nella seconda colonna di destra invece, si trovano i link agli ultimi commenti inseriti nei vari articoli del blog e, inoltre, una serie di link utili per migliorare la ricerca dell'utente, tra cui i siti dei principali APT (Agenzia Per il Turismo) della Toscana, il sito ufficiale della Regione e del Turismo e quello del Consorzio Lamma per le previsioni meteorologiche. Per aumentare la popolarità del sito, si è lavorato molto per migliorare il linking interno, inserendo negli articoli, dove possibile, link che rimandassero sia ai vari tag sia ad altri articoli correlati. A volte, sono stati creati volutamente anche articoli di appoggio per pubblicizzare altri articoli. Inoltre, si è



## 8.4 Analisi delle statistiche

L'analisi delle statistiche del sito *toscanaper.me* è stata effettuata mediante lo strumento Google Analytics. Come mostrato in figura 9.5, la dashboard di Google Analytics ci fornisce una panoramica dei dati principali relativi ai primi tre mesi di vita del sito.

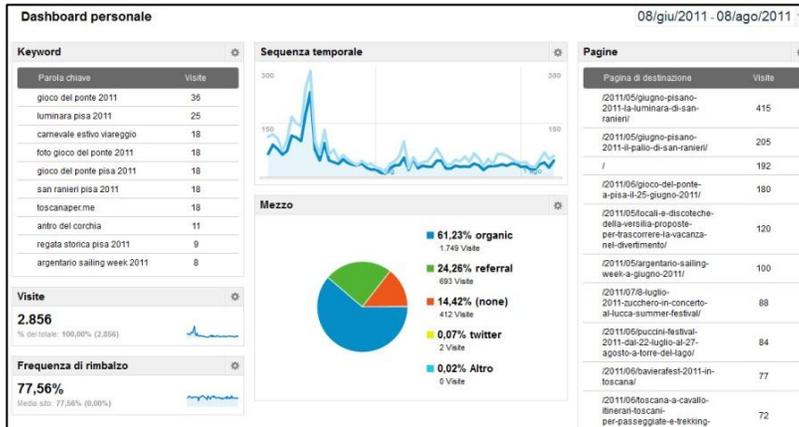


Figura 9.5 Dashboard di *toscanaper.me* su Google Analytics

Nei mesi presi in analisi, nello specifico dall'8 Giugno all'8 Agosto 2011, *toscanaper.me* ha avuto complessivamente **2.856 visite totali**. Di queste, l'**87,4%** è dato da nuovi visitatori e il restante **12,29%** da visitatori di ritorno. In ottica di visibilità e awareness del sito questi dati sono quindi soddisfacenti, mentre lo sono meno dal punto di vista della fidelizzazione dell'utente. La media delle pagine visualizzate per visita è di **1,44**, con un tempo medio sul sito di circa 1 minuto e 10 secondi e una frequenza di rimbalzo pari al **77,56%**. Nonostante questi dati dimostrino una tipologia di visita al sito breve e poco approfondita, tipica dell'utente che cerca solo informazioni mirate, l'aspetto positivo è che più delle metà degli accessi, il **61,23%**, proviene da risultati organici dei motori di ricerca, a indicare che i contenuti del sito sono ben indicizzati. A seguire, nel grafico troviamo un **24,26%** di visite provenienti da siti referenti, che indicano un livello di popolarità abbastanza positivo, e un **14,42%** di accessi diretti.

Per quanto riguarda l'analisi dei contenuti, le principali keyword con cui i visitatori trovano il sito, corrispondono per lo più a termini riferiti a eventi. Tra questi, troviamo il Giugno pisano, il Festival Pucciniano, l'Argentario Sailing Week, il concerto di Zucchero al Lucca Summer Festival e il BavieraFest, tutti eventi molto conosciuti a

livello nazionale. Infatti, come si vede in figura 9.6, delle prime dieci pagine di destinazione principale, sette appartengono proprio alla categoria degli eventi sopra citati, mentre gli altri riguardano per lo più le attività o gli itinerari.

Pagina di destinazione	Visite	Rimbalti	Frequenza di rimbalzo
1. /2011/05/giugno-pisano-2011-la-luminara-di-san-raineri/	415	348	83,86%
2. /2011/05/giugno-pisano-2011-il-pallo-di-san-raineri/	205	184	89,76%
3. /	192	93	48,44%
4. /2011/05/gioco-del-ponte-a-pisa-il-25-giugno-2011/	180	163	90,56%
5. /2011/05/foccali-e-discoche-della-versilia-proposte-per-trascorrere-la-vacanza-nei-divertimento/	120	89	74,17%
6. /2011/05/targetario-sailing-week-a-giugno-2011/	100	84	84,00%
7. /2011/07/8-luglio-2011-zucchero-in-concerto-al-fuoca-summer-festival/	88	79	89,77%
8. /2011/06/puccini-festival-2011-dal-22-luglio-al-27-agosto-a-torre-del-lago/	84	56	66,67%
9. /2011/06/baiafest-2011-in-toscana/	77	69	89,61%
10. /2011/06/toscana-a-cavallo-itinerari-toscani-per-passeggiate-e-trekking-a-cavallo/	72	50	69,44%

Figura 9.6 Pagine di destinazione principali di *toscanaper.me*

Sintomatici, come si vede in figura 9.7, sono i dati relativi agli accessi tramite Facebook grazie alla condivisione dei post, a conferma di quanto sia importante pianificare una strategia anche sui social network. **252** è il numero delle visite totali, che corrispondono all'8,82% delle visite totali del sito. La media di pagine per visita è di **1,90** (il 32,65% in più rispetto alla media del sito), il tempo medio è di 5 minuti e 30 secondi (il 380,29% in più della media del sito) e la frequenza di rimbalzo è del **70,63%** (l'8,92% in meno della media del sito).

Percorso referral	Pagina di destinazione	Visite	Pagine/visita	Tempo medio sul sito	% nuove visite	Frequenza di rimbalzo
1. /A.php	/2011/05/gioco-del-ponte-a-pisa-il-25-giugno-2011/	39	1,23	00:00:10	84,62%	97,44%
2. /A.php	/	36	3,08	00:08:16	11,11%	25,00%
3. /A.php	/2011/05/social-tools-4-tourism-a-pisa/	21	1,38	00:05:55	71,43%	76,19%
4. /A.php	/2011/05/la-passeggiata-filosofica-alla-notte-blu-della-costa-degli-etruschi/	19	1,58	00:01:05	78,95%	73,68%
5. /A.php	/2011/06/baiafest-2011-in-toscana/	17	1,18	00:00:23	82,35%	82,35%
6. /A.php	/2011/05/giugno-pisano-2011-la-luminara-di-san-raineri/	16	1,50	00:00:22	81,25%	87,50%
7. /A.php	/2011/06/pisa-blog-tour-2011-dal-15-al-17-giugno/	11	1,09	00:00:04	54,55%	90,91%
8. /A.php	/2011/07/effetto-venezia-2011-a-ivorno/	10	1,20	00:00:16	60,00%	90,00%
9. /A.php	/2011/05/pisa-gioca-11-e-12-giugno-2011/	8	1,00	00:00:00	87,50%	100,00%
10. /A.php	/2011/07/la-festa-dell'innamorata-il-14-e-15-luglio-2011-a-capovivri/	7	1,29	00:00:05	28,57%	85,71%

Figura 9.7 Pagine di destinazione principali di *toscanaper.me* tramite Facebook

Le pagine di destinazione principale, appartengono ancora alla categoria degli eventi, ma a differenza delle pagine di destinazione provenienti dai motori di ricerca, queste corrispondono a eventi di interesse prettamente giovanile, che sono state condivise più volte su Facebook direttamente dal sito. Questo è un aspetto di valutazione importante. Considerato che il target preferenziale di *toscanaper.me* è quello giovanile, e che proprio questo target dimostra interesse verso argomenti rivolti ai giovani, possiamo ritenere positivo il lavoro svolto nella scelta degli argomenti e nella scrittura degli articoli, e confermare l'importanza dei social network per veicolare le informazioni e aumentare la popolarità.

## 9. Conclusioni

L'obiettivo che si prefigge questa tesi, è quello di illustrare le strategie in grado di migliorare il posizionamento di un sito nelle SERP dei motori di ricerca, analizzando i principali fattori SEO on-page e off-page. Questi metodi, sono stati applicati al sito *toscanaper.me*, in modo da favorire la visibilità e migliorare il posizionamento di questo nuovo sito.

Sulla base dei dati raccolti e analizzati con Google Analytics, per quanto riguarda la visibilità e il posizionamento del sito sui motori di ricerca, si può dire di aver raggiunto l'obiettivo: il numero delle visite è aumentato nel tempo e le pagine risultano essere ben indicizzate dai motori. Per quanto riguarda invece la fidelizzazione dell'utente, è opportuno pianificare una strategia per aumentare la percentuale di visitatori di ritorno. Inoltre, il tempo medio sul sito è basso e la percentuale della frequenza di rimbalzo relativamente alta. Questo comportamento dell'utente non è comunque da considerarsi del tutto negativo, perché riflette implicitamente lo scopo del sito: essere trovati attraverso una ricerca di parole chiave di nicchia, che restituisca pagine in grado di fornire informazioni utili e immediate, cosa ben dimostrata dalle visite relative alle pagine sugli eventi. In questi casi, infatti, è molto frequente che l'utente, una volta visitata la pagina, esca dal sito soddisfatto delle informazioni che ha trovato, oppure, in caso contrario, esca per continuare la propria ricerca altrove. In ogni caso, è necessario lavorare ancora per aumentare la popolarità e veicolare meglio le informazioni. Fondamentale sarebbe pubblicare costantemente nuovi articoli e condividerli sui principali social network, con una frequenza ottimale di due volte a settimana. Un sito come questo, infatti, per non perdere visibilità, ha bisogno di essere sempre aggiornato, a maggior ragione se il sito in questione è nuovo, con poca link popularity, e una vasta concorrenza.

Concludendo, porgo le mie considerazioni alla luce del lavoro svolto: ogni sito è diverso dall'altro, di conseguenza, ogni sito necessita una strategia mirata che risponda alla tipologia di target a cui si vuole rivolgere e agli obiettivi che vuole raggiungere. In ogni caso, conviene implementare le tecniche SEO, e dedicare una cura particolare alla

qualità dei contenuti. Con pochi sforzi, il sito può fare la differenza e raggiungere risultati soddisfacenti, facendo emergere anche un sito nuovo in settori molto competitivi.

## Appendice A

### PAROLE CHIAVE PER LA CATEGORIA EVENTI

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[eventi in toscana]	0,23	880	880
[eventi e sagre in toscana]	0,22	140	140
[eventi in toscana oggi]	0,31	210	210
[mercatini in toscana]	0,16	210	210
[mercatini di natale in toscana]	0,32	260	260
[mercatini antiquariato in toscana]	0,25	36	36
[manifestazioni Toscana]	0,19	110	110
[lucca summer festival]	0,11	2900	1900
[summer festival lucca 2011]	0,11	1600	1300
[lucca summer festival 2011 programma]	0,1	260	260
[festival puccini]	0,07	880	590
[torre del lago puccini festival]	0,09	110	91
[festival puccini 2011]	0,09	260	210
[versiliana festival]	0,01	58	58

### PAROLE CHIAVE PER LA CATEGORIA ITINERARI

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[toscana itinerari turistici]	0,53	58	58
[itinerari turistici toscana]	0,48	140	140
[toscana itinerari]	0,32	320	320
[itinerari moto toscana]	0,42	260	260
[itinerari mtb toscana]	0,24	36	36
[itinerari enogastronomici]	0,45	140	110
[itinerari enogastronomici toscana]	0,55	36	36
[escursioni a cavallo]	0,54	210	210
[escursioni a cavallo toscana]	0,5	36	36

[itinerari in bicicletta toscana]	0,44	16	16
[visitare la toscana]	0,41	320	260
[cosa visitare a firenze]	0,23	880	880
[posti da visitare a firenze]	0,32	110	110
[luoghi da visitare a firenze]	0,23	110	110

### PAROLE CHIAVE PER LA CATEGORIA ATTIVITA'

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[trekking toscana]	0,27	480	480
[trekking in toscana]	0,28	91	91
[toscana itinerari]	0,32	320	320
[percorsi trekking toscana]	0,28	46	46
[itinerari mtb toscana]	0,24	36	36
[camping in toscana]	0,55	320	260
[camping village in toscana]	0,43	58	46
[escursioni a cavallo]	0,54	210	210
[escursioni a cavallo toscana]	0,5	36	36
[itinerari in bicicletta toscana]	0,44	16	16
[noleggio barche toscana]	0,83	170	170
[noleggio barche a vela toscana]	0,9	260	260
[campi da golf in toscana]	0,35	28	28
[golf in toscana]	0,37	140	110

### PAROLE CHIAVE PER LA CATEGORIA CULTURA

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[mostre in toscana]	0,52	110	110
[arte in toscana]	0,18	36	36
[toscana itinerari]	0,32	320	320
[cultura toscana]	0,1	91	73

[settimana della cultura toscana]	0,08	16	16
[musei in toscana]	0,14	46	46
[musei di firenze]	0,47	390	320
[elenco musei di firenze]	0,31	36	36
[città etrusche]	0,01	210	210
[città etrusche in toscana]	0,06	16	16
[monumenti di firenze]	0,11	390	390
[i monumenti di firenze]	0,08	73	58

### PAROLE CHIAVE PER LA CATEGORIA TERRITORIO

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[chianti toscana]	0,27	590	140
[toscana chianti]	0,57	260	91
[zona chianti toscana]	0,32	22	22
[val d'orcia toscana]	0,61	170	170
[toscana val d orcia]	0,49	46	46
[isole toscana]	0,21	260	260
[isole della toscana]	0,14	170	170
[alpi apuane]	0	2900	2900
[parco alpi apuane]	0	210	210
[alpi apuane toscana]	0,02	22	22
[terme in toscana]	0,08	4400	4400
[le terme in toscana]	0,7	46	46
[maremma toscana]	0,66	2900	2400
[parco della maremma toscana]	0,66	73	73

### PAROLE CHIAVE PER LA CATEGORIA ENOGASTRONOMIA

<b>Parola chiave</b>	<b>Concorrenza tra inserzionisti</b>	<b>Volume di ricerca locale</b>	<b>Volume di ricerca globale</b>
[enogastronomia toscana]	0,33	58	58
[itinerari enogastronomici]	0,45	140	110

[itinerari enogastronomici toscana]	0,55	36	36
[montalcino]	0,14	12100	5400
[brunello di montalcino]	0,58	12100	2900
[rosso di montalcino]	0,31	1900	390
[cantine montalcino]	0,14	91	73
[cucina toscana]	53	2400	1000
[la cucina toscana]	0,21	110	58
[ricette cucina toscana]	0,55	170	158
[degustazioni toscana]	0,28	36	36
[degustazioni vino toscana]	0,34	16	16

## Riferimenti bibliografici

[1] John I. Jerkovic. *SEO. L'arte dell'ottimizzazione per i motori di ricerca*. Milano, Tecniche Nuove, collana Hops Technologie, 2010.

[2] Giorgio Taverniti. *SEO power*. Serie a cura di Luca Conti. Milano, Hoepli, 2011.

[3] Abhishek Das, Ankit Jain, *Indexing the World Wide Web: The Journey So Far, Next Generation Search Engines: Advanced Models for Information Retrieval*, 2011

[[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/research.google.com/it/pubs/archive/37043.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/it/pubs/archive/37043.pdf)]

[4] Stefan Buettcher, Charles L. A. Clarke, Gordon V. Cormack, *Information Retrieval: Implementing and Evaluating Search Engines*, 2010

[<http://www.ir.uwaterloo.ca/book/01-introduction.pdf>]

[5] Page L. *Method for node ranking in a linked database*, United State Patent, 1998

[<http://patft.uspto.gov/netacgi/nph->

[Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fmetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=6285999.PN.&OS=PN/6285999&RS=PN/6285999](http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fmetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=6285999.PN.&OS=PN/6285999&RS=PN/6285999)]

[6] Sahami M., Heilman, T. D., *Generating query suggestions using contextual information*, United State Patent, 2011

[<http://patft.uspto.gov/netacgi/nph->

[Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fmetahtml%2FPTO%2Fsearch-bool.html&r=10&f=G&l=50&co1=AND&d=PTXT&s1=google&OS=google&RS=google](http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fmetahtml%2FPTO%2Fsearch-bool.html&r=10&f=G&l=50&co1=AND&d=PTXT&s1=google&OS=google&RS=google)]

[7] Gyöngyi Z., Garcia-Molina H., Pedersen J., *Combating Web Spam with TrustRank*

[<http://www.vldb.org/conf/2004/RS15P3.PDF>]

[8] Google Research

[<http://research.google.com/>]

[9] *Guida di Google all'ottimizzazione per i motori di ricerca*

[[http://static.googleusercontent.com/external\\_content/untrusted\\_dlcp/www.google.com/it/intl/it/webmasters/docs/search-engine-optimization-starter-guide-it.pdf](http://static.googleusercontent.com/external_content/untrusted_dlcp/www.google.com/it/intl/it/webmasters/docs/search-engine-optimization-starter-guide-it.pdf)]

[10] *Guida di Google Strumenti per i Webmaster*

[<http://www.google.com/support/webmasters/bin/answer.py?hl=it&answer=35291>]

[11] *Guida di Yahoo! Search Content Quality Guidelines*

[<http://help.yahoo.com/l/us/yahoo/search/basics/basics-18.html?pir=ahs0yK1ibUlt9V4vBr3XrhhpOnm5d4MMT6idh9lndPdz8wkGTrI->]

[12] Yahoo! Research

[<http://research.yahoo.com/>]

[13] Bing Webmaster Central Blog

[[http://www.bing.com/community/site\\_blogs/b/webmaster/archive/2009/09/03/search-engine-optimization-for-bing.aspx](http://www.bing.com/community/site_blogs/b/webmaster/archive/2009/09/03/search-engine-optimization-for-bing.aspx)]

[14] *Bing Webmaster Tools Table of Contents*

[<http://onlinehelp.microsoft.com/en-us/bing/hh204434.aspx>]

[15] Search Engine Watch, *Successful SEO Tactics: On-site Optimisation*

[<http://searchenginewatch.com/article/2101314/Successful-SEO-Tactics-On-Site-Optimization>]

[16] WordPress.org

[<http://wordpress.org/>]

[17] WordPress.org, *WordPress Plugin*

[<http://wordpress.org/extend/plugins/>]