

**UNIVERSITÀ di PISA**



Corso di Laurea in  
**INFORMATICA UMANISTICA**



InformaticaUmanistica

*“Valutazione dell’efficacia di  
tecniche di web authoring per il  
miglioramento del ranking nei  
motori di ricerca”*

Tesi di Laurea di  
Azzurra Toracca

Relatore:  
Professor Vincenzo Gervasi

Anno Accademico 2007/2008

*Un ringraziamento particolare ad Ariatel Web s.r.l. per la  
disponibilità dimostrata durante la realizzazione della tesi di  
laurea*

# INDICE

<b>Introduzione .....</b>	<b>5</b>
<b>1 Le tecniche SEO .....</b>	<b>8</b>
<b>1.1 Definizione di tecniche SEO .....</b>	<b>8</b>
<b>1.2 La struttura dei dati.....</b>	<b>8</b>
1.2.1 Il linguaggio HTML .....	9
1.2.2 I metadati.....	10
<b>1.3 Le caratteristiche dei dati.....</b>	<b>14</b>
1.3.1 Keywords Density .....	14
1.3.2 BackLink .....	14
<b>1.4 Le sitemap.....</b>	<b>15</b>
<b>2 Realizzazione del sito web “Indicizzazione e posizionamento nei motori di ricerca” .....</b>	<b>16</b>
<b>2.1 Analisi di blog e siti web.....</b>	<b>16</b>
<b>2.2 Creazione della scaletta argomentativa .....</b>	<b>17</b>
<b>2.3 Realizzazione della grafica.....</b>	<b>19</b>
<b>2.4 Creazione del sito web .....</b>	<b>20</b>
<b>2.5 Creazione del dominio e pubblicazione in rete.....</b>	<b>21</b>
<b>2.6 Segnalazione dei siti web ai motori di ricerca.....</b>	<b>23</b>
2.6.1 Strumenti di Google per i webmaster .....	23
2.6.2 Yahoo!: proponi l’inclusione del tuo sito .....	25
2.6.3 Segnala il tuo sito a Live Search .....	26
2.6.4 Il MioAsk per l’inserimento di pagine web .....	27
2.6.5 Motori di ricerca Virgilio e Arianna.....	28
<b>3. Realizzazione del sito web “Indicizzazione e posizionamento nei motori di ricerca” .....</b>	<b>29</b>
<b>3.1 Scelta delle parole chiave .....</b>	<b>29</b>
<b>3.2 Metodi di ottimizzazione utilizzati.....</b>	<b>30</b>
<b>3.3 Modifica della struttura dei domini ai fini dell’indicizzazione .....</b>	<b>32</b>
3.3.1 Uso dei meta tag ai fini dell’indicizzazione.....	32
3.3.2 Uso dei link d’entrata e d’uscita da una pagina web .....	34
3.3.3 Creazione di una sitemap .....	44
3.3.4 Uso dei tag HTML per l’ottimizzazione .....	46
3.3.5 Sito web non modificato da alcun metodo di ottimizzazione .....	47
3.3.6 Sito web ottimizzato attraverso l’uso delle tecniche SEO inserite separatamente negli altri siti web .....	47
<b>4. Raccolta e analisi dei dati conclusivi.....</b>	<b>49</b>
<b>4.1 Informazioni statistiche ricavate da stat.ariatel.it .....</b>	<b>49</b>
4.1.1 Dati statistici suddivisi per domini web .....	52
4.1.2 Sintesi e comparazione dei dati statistici .....	66
<b>4.2 Raccolta e analisi dei dati relativi al posizionamento.....</b>	<b>69</b>
4.2.1 Google .....	69
4.2.2 Comparazione dei dati ottenuti per il motore di ricerca Google .....	80
4.2.3 Ask.com .....	82
4.2.4 Live Search e Yahoo! .....	83

4.2.1 Arianna e Virgilio.....	84
<b>5. Conclusioni .....</b>	<b>85</b>
<b>5.1 Comportamento dei motori di ricerca .....</b>	<b>85</b>
<b>5.2 Effetti delle tecniche di ottimizzazione .....</b>	<b>87</b>
<b>Webliografia e Sitografia .....</b>	<b>89</b>

## INTRODUZIONE

I motori di ricerca permettono di inserire in un'apposita maschera presente nell'interfaccia grafica parole o frasi per specificare l'argomento che i documenti restituiti dovranno trattare. Una volta avviato il processo di ricerca, le possibili risposte alla query d'entrata vengono ricercate all'interno del database del motore stesso. Quest'ultimo quindi restituisce i risultati in una lista ordinata per coerenza di risultati e popolarità delle pagine; tale lista è detta SERP (acronimo di "Search Engine Report Page", "pagina dei risultati di un motore di ricerca").

La creazione del database avviene ad opera del lavoro di robot e spider che visitano le pagine presenti in rete e associano a queste un particolare valore che ne determina la posizione all'interno della SERP. Robot e spider non lavorano autonomamente, ma seguono le regole dettate da uno specifico algoritmo di ricerca, utilizzato da quel particolare crawler.

Per fare quindi in modo che un sito web sia indicizzato da un motore di ricerca, occorre strutturarne secondo le specifiche che l'algoritmo utilizzato dal crawler valuta per assegnare un valore più o meno alto ad una specifica pagina.

Il problema della visibilità di un sito web è fondamentale per tutti coloro che creano un sito e che auspicano di ricevere il maggior numero possibile di visite.

Studiando il comportamento di robot e spider e conoscendo gli algoritmi di alcuni crawler, si sono sviluppate negli anni alcune tecniche che permettono di ottimizzare il proprio sito in funzione del posizionamento all'interno di un motore di ricerca. Tali tecniche sono generalmente indicate con l'acronimo SEO, "Search Engine Optimization" (letteralmente "Ottimizzazione per i motori di ricerca").

Scopo della presente tesi di laurea è quello di affrontare i due problemi illustrati brevemente, ovvero il problema della ricerca e il problema della visibilità di un sito web, e di verificare l'effettiva efficacia di alcune tra le tecniche di ottimizzazione maggiormente utilizzate.

Sono stati quindi preventivamente visitati una serie di siti web e blog che trattano l'argomento interessato, in modo da ottenere informazioni utili a due differenti scopi: scegliere con giudizio le tecniche SEO da analizzare e applicare durante il lavoro di analisi; decidere gli argomenti da trattare nel sito web realizzato allo scopo di testare direttamente le tecniche prese in considerazione.

Completata l'analisi, proceduto con la valutazione delle possibili strategie da adottare, sono state scelte quattro tecniche di ottimizzazione. Queste sono state successivamente applicate ad un pari numero di domini web pubblicati in rete, il cui contenuto e la cui struttura risultano assolutamente identici tra loro. Questa strategia ha permesso di procedere ad una comparazione dei risultati ottenuti in fase conclusiva.

È stato inoltre deciso di creare due ulteriori domini: uno rimasto del tutto invariato, senza l'applicazione di alcuna tecnica; l'altro invece a cui sono state applicate tutte le tecniche viste separatamente negli altri domini.

Completata la fase di realizzazione del sito, i differenti domini sono stati sottoposti all'attenzione di una serie di motori di ricerca, attraverso gli strumenti messi a disposizione da ognuno di essi. La scelta è ricaduta essenzialmente su sei crawler: Google, Yahoo!, Live Search, Ask.com, Virgilio, Arianna. Tale scelta è stata dettata dal desiderio di analizzare il funzionamento sia dei maggiori crawler, utilizzati a livello mondiale (quindi Google, Yahoo!, Live Search e Ask.com), sia di quei motori di ricerca minori, il cui utilizzo si restringe al territorio nazionale (Arianna e Virgilio), e poter quindi comparare l'efficacia che le tecniche SEO adottate dimostrano nei due diversi ambiti. Sfortunatamente, come sarà illustrato nel capitolo 4, i risultati ottenuti per i motori di ri-

cerca italiani non sono stati significativi a causa della loro stretta dipendenza dal motore di ricerca statunitense Google.

Trascorse due settimane dalla pubblicazione in rete e decise le parole chiave su cui si desiderava eseguire l'ottimizzazione, attraverso l'utilizzo del programma Accurate Monitor for Search Engine<sup>1</sup>, sono state effettuate delle scansioni per ogni sito web, allo scopo di controllare il posizionamento di questi ultimi su ogni motore di ricerca preso in considerazione.

Creato il sito web, trascorse le due settimane ed eseguite le scansioni, si è passati alle modifiche dei domini attraverso l'uso delle tecniche scelte in precedenza: uso dei meta tag, linguaggio HTML, Sitemap e uso dei BackLink. Tali tecniche sono tutte rigorosamente gratuite, in primo luogo perché l'utilizzo di diverse tecniche a pagamento avrebbe costituito un costo decisamente alto da sostenere per un progetto di ricerca, che non ha come scopo un eventuale guadagno; secondariamente per dimostrare che anche l'utilizzo di piccoli accorgimenti, del tutto gratuiti, può aiutare a ottimizzare il sito web creato.

Pubblicate le nuove modifiche, si sono dovute attendere tre settimane, per far sì che motori di ricerca visitassero nuovamente i nostri siti web e fare in modo che questi migliorassero il loro posizionamento.

Trascorso quest'intervallo di tempo, sono state effettuate nuove scansioni, per ottenere i dati definitivi della nostra ricerca, e sono state analizzate le statistiche riguardanti i singoli domini. Completata la fase di raccolta, si è proceduto con l'analisi e con la comparazione dei dati per stilare le conclusioni del caso.

---

<sup>1</sup> Accurate Monitor for Search Engine: prodotto della CleverStat che permette di monitorare la posizione di un sito web su i motori di ricerca più conosciuti per una serie di parole chiave o frasi. Ne esistono due versioni: a pagamento e gratuita per 30 giorni di prova, quest'ultima con alcune funzionalità ridotte. Il programma è scaricabile all'indirizzo [www.cleverstat.com/en/search-engines.htm](http://www.cleverstat.com/en/search-engines.htm). Ai fini di questa ricerca, è stata usata la versione a pagamento messa a disposizione dall'azienda Ariatel Web s.r.l.

## **1. LE TECNICHE SEO**

In questo capitolo saranno forniti accenni teorici utili per comprendere al meglio l'analisi effettuata e i risultati ottenuti.

Sarà quindi chiarito il termine SEO e verranno illustrate nel dettaglio le tecniche utilizzate nel corso della sperimentazione: il codice HTML, l'uso di metadati e meta tag, i BackLink, le Sitemap.

### **1.1 DEFINIZIONE DI TECNICHE SEO**

Con il termine SEO, acronimo di Search Engine Optimization, si indica l'insieme di tutte le attività utili a migliorare il posizionamento di un sito o di una pagina web all'interno della SERP di un motore di ricerca.

Si tratta di attività complesse di strutturazione dei contenuti delle pagine web di un sito, che fanno parte di una vasta attività di SEM (Search Engine Marketing), il marketing dei motori di ricerca.

Queste tecniche quindi interessano molteplici fattori: la struttura dei dati e i contenuti della pagina stessa; l'insieme di metadati e meta tag; le caratteristiche della pagina web creata, come la densità delle parole chiave e il numero e l'importanza dei cosiddetti BackLink.

### **1.2 LA STRUTTURA DEI DATI**

La struttura dei dati è un elemento molto importante per l'indicizzazione di una pagina web. I motori di ricerca infatti leggono il contenuto della pagina per determinare il tema trattato e le parole, o insiemi di parole, da considerare determinanti per la pagina stessa.

Per migliorare quest'aspetto è possibile andare a lavorare sia sulla struttura della pagina web, quindi sul linguaggio HTML che costituisce la base di ogni sito

internet, sia sulla parte che costituisce i cosiddetti metadati, che interessano meta tag e attributo alt sulle immagini.

### 1.2.1 IL LINGUAGGIO HTML

Creare un sito web ottimizzato nella “pulizia” del codice HTML può rivelarsi utile per ottenere buone posizioni sui motori di ricerca.

I crawler infatti troveranno meno difficoltà nell'interpretare un codice chiaro ed essenziale e, in alcuni casi, potrebbero assegnare un punteggio maggiore per quei siti web validati secondo le direttive W3C<sup>2</sup>.

È quindi fondamentale prestare molta attenzione alla corretta scrittura del codice ed evitare di ripetere tag “unici”, come per esempio il tag title.

Inoltre, i motori di ricerca, Google in primis, assegnano maggiore o minore importanza agli elementi testuali presenti nella pagina web in base ai tag che li contengono.

Ciò significa che verrà assegnata maggiore importanza alle parole che compaiono all'interno del tag `<h1></h1>` rispetto a quelle presenti in tag `<h2></h2>`; a sua volta queste ultime saranno considerate maggiormente rispetto a parole contenute nel tag `<p></p>`.

Ugualmente importanti sono da considerare i tag di formattazione del testo (bold, italic, underline) e di citazione (tag `<em></em>`).

Naturalmente, come per tutti i metodi di ottimizzazione, è importante ricordare che prima di tutto bisogna dare importanza all'usabilità della pagina web. Fondamentale è quindi evitare di abusare di questi metodi e non forzare il testo con l'inserimento di una quantità eccessiva di tag di formattazione.

---

<sup>2</sup> W3C: acronimo di World Wide Web Consortium. Fondato nel 1994 da Tim Berners Lee, padre del Web, ha lo scopo di migliorare i protocolli e i linguaggi per il World Wide Web e per aiutare il Web a migliorare le sue potenzialità. Sito ufficiale: [www.w3.org](http://www.w3.org)

### 1.2.2 I METADATI

Come concetto generale, con metadati s'intendono quegli elementi che indicano proprietà, caratteristiche e significato dei dati presenti in un determinato archivio e che li relazionano. I metadati quindi sono quegli elementi che permettono di conoscere le caratteristiche descrittive di ciò che l'archivio contiene.

Per quanto riguarda la programmazione web, con il termine metadati s'intendono tutti quei dati cosiddetti sensibili inclusi nei tag meta, ma che in ogni caso esplicano la funzione sopra descritta, cioè quella di riepilogare il contenuto della pagina stessa.

I metadati rilevanti per l'attività di ottimizzazione di una pagina web sono i meta tag sopra menzionati e l'attributo alt da associare alle immagini inserite nella pagina web.

#### *Alt sulle immagini*

L'alt sulle immagini è un attributo molto importante da tenere in considerazione, prima di tutto perché è esplicitamente richiesto dalle direttive del W3C in termini di usabilità di una pagina web, ma anche perché può risultare molto utile per l'attività di ottimizzazione della pagina stessa.

È importante inserire quest'attributo all'interno del tag `<img />`, ma fondamentale è anche la scelta delle parole utilizzate per la descrizione dell'immagine.

Nel caso in cui questa funzioni da link, la sua descrizione deve richiamare il contenuto della pagina linkata; in caso contrario è molto importante scegliere parole che descrivano brevemente ciò che l'immagine mostra. Ai fini dell'ottimizzazione, se è possibile, è utile inserire una parola chiave all'interno dell'attributo.

Resta comunque fondamentale evitare di inflazionare l'uso di parole chiave, per evitare il rischio di spam e quindi della mancata indicizzazione da parte dei motori di ricerca.

### *Meta tag*

I meta tag sono “etichette” proprie del linguaggio HTML utilizzate per aggiungere meta-informazioni alla pagina web realizzata.

Scopo di queste etichette è quello di fornire informazioni aggiuntive ai motori di ricerca ai fini dell'ottimizzazione del sito web; a differenza degli altri tag del linguaggio HTML, i meta non modificano in alcun modo la formattazione della pagina.

Esistono molti meta tag differenti, alcuni con funzioni molto specifiche (per esempio i meta esclusivi per Google), ma se ne possono considerare essenzialmente sei indispensabili: meta title, meta description, meta keywords, meta robots, meta pragma, meta revisit-after.

- Meta title: meta tag fondamentale, per il quale il W3C raccomanda la differenziazione per ogni pagina che compone il sito web in esame. L'ottimizzazione di questo meta è molto utile ai fini dell'indicizzazione del sito; è quindi fondamentale creare un titolo che rispecchi al meglio il contenuto della pagina stessa e che contenga le parole-chiave per cui vogliamo indicizzare quella stessa pagina. Contemporaneamente, è opportuno cercare di non rendere il nostro titolo troppo lungo poiché i motori di ricerca visualizzano mediamente 50 caratteri.

Questa è la sintassi del tag TITLE

```
<title> Titolo della pagina </title>
```

Esiste anche la dicitura

```
<meta name="title" content="Titolo della pagina">
```

ma lo standard W3C consiglia l'uso della prima sintassi.

- Meta description: questo meta illustra una breve descrizione dei contenuti della pagina in cui è collocato; molti motori di ricerca visualizzano tale descrizione accanto al titolo della pagina o al posto di essa nella lista dei risultati ottenuti.

Anche in questo caso il testo non deve essere eccessivamente lungo; bisogna inoltre inserire una descrizione discorsiva ma semplice, senza produrre una lista di parole chiave.

La sintassi del meta tag DESCRIPTION è la seguente:

```
<meta name="description" content="Descrizione del contenuto della pagina">
```

- Meta keywords: il contenuto di questo meta è costituito da una lista di parole chiave attraverso le quali desideriamo indicizzare una pagina web. La scelta di tale lista è molto importante poiché è fondamentale che le parole inserite in questo tag compaiano all'interno del testo della pagina.

Nella fase di scelta delle parole chiave, può essere utile calarsi nei panni dell'utente e cercare di immaginare quali parole potrebbe utilizzare per la sua ricerca; è perciò importante tener conto che le persone possono utilizzare termini diversi per ottenere i medesimi risultati.

La sintassi del tag KEYWORDS è:

```
<meta name="keywords" content="parola1, parola2, ..., parolaN">
```

- Meta robots: diversamente dai tag title, description e keywords, il cui scopo è quello di includere informazioni supplementari alla pagina, il meta ROBOTS contiene dei "consigli di comportamento" per gli spider che hanno lo scopo di prelevare e analizzare la pagina.

Attraverso il meta ROBOTS è possibile impartire due diverse direttive agli spider:

- chiedere di indicizzare (INDEX) o meno (NOINDEX) i contenuti della pagina visitata;
- chiedere di seguire tutti i link contenuti nella pagina (FOLLOW) o di non seguirli (NOFOLLOW).

Questi parametri possono essere combinati a piacimento all'interno del content del tag meta.

La sintassi è quindi la seguente:

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

- Meta pragma con contenuto no-cache: scopo di questo meta tag è quello di indicare al browser se questo può effettuare o meno un salvataggio dei contenuti della pagina nella memoria cache del computer. Questa la sintassi del meta PRAGMA:

```
<meta http-equiv="pragma" content="no-cache">
```

In caso di pagine in continuo aggiornamento si consiglia l'attivazione del no-cache in modo da avere maggiore controllo sui contenuti.

- Meta revisit-after: ha lo scopo di richiedere allo spider di tornare a visitare la pagina ad intervalli regolari, il cui tempo è stabilito nel content

```
<meta name="Revisit-after" content="210 days">
```

## **1.3 LE CARATTERISTICHE DEI DATI**

### **1.3.1 KEYWORDS DENSITY**

Parlare di Keywords Density, la densità delle parole-chiave, significa affrontare un argomento su cui gli esperti di indicizzazione dibattono da molto tempo.

Il calcolo della densità delle parole chiave non è altro che un indicatore di quante volte una lista di parole compare nella pagina web analizzata; a questo scopo esistono molti programmi che, data in entrata l'URL della pagina stessa, ne restituiscono un'analisi in termini numerici delle parole rilevanti che vi compaiono (non sono quindi prese in considerazione parole di uso comune come congiunzioni e articoli).

Bisogna dire che non esiste una densità di parole-chiave ideale, da usare come modello a cui paragonare i risultati ottenuti dall'analisi; ogni pagina web infatti ha la propria densità ideale che fa sì che il testo prodotto comunichi in modo chiaro, semplice e diretto il contenuto stesso.

Recentemente sempre più esperti del settore affermano che la keywords density non è un parametro determinante per il buon posizionamento o meno all'interno di una lista di risultati di un motore di ricerca.

L'unico parametro a cui invitano di prestare attenzione è quello di non esagerare e non sovraccaricare eccessivamente la pagina web di parole-chiave, soprattutto se posizionate in maniera non consona al contenuto della pagina, per evitare il conseguente spamming da parte dei motori di ricerca.

### **1.3.2 BACKLINK**

Con il termine BackLink s'intende il contatto che una pagina web riceve da un'altra; qualità e valori di questi contatti generano la cosiddetta Link Popularity del sito.

I BackLink sono un fattore molto importante che permette di influenzare positivamente la posizione di una pagina web all'interno della lista di risultati fornita da un motore di ricerca.

Naturalmente ciò che conta veramente in questo tipo di parametro non è soltanto il numero di link ricevuti ma, soprattutto, la loro qualità: meglio quindi ricevere pochi link da siti autorevoli piuttosto che molti link da siti minori.

Per far in modo che un sito, soprattutto a livello amatoriale, sia indicizzato, o quanto meno conosciuto dalla comunità virtuale, si stanno moltiplicando siti e blog in cui si offrono scambi di link.

#### **1.4 LE SITEMAP**

Esistono essenzialmente due tipi differenti di Sitemap: una sitemap di tipo testuale, inserita all'interno di una pagina web, e una sitemap in formato XML<sup>3</sup>.

La sitemap di tipo testuale elenca gerarchicamente tutte le pagine presenti nel sito web, con i relativi link alle stesse. Oltre a rivestire una discreta importanza nell'attività di scansione della rete da parte degli spider dei motori di ricerca, l'inserimento di una sitemap è molto utile per facilitare la navigazione dell'utente all'interno del sito.

La sitemap in formato XML è stata introdotta da Google per facilitare la scansione delle pagine web da parte dei propri spider. Anche questo tipo di mappa contiene una lista di link, un elenco di tutte le URL delle pagine di un sito web.

Recentemente Google ha messo a disposizione dei webmaster una serie aggiuntiva di sitemap: sitemap per Google News, sitemap per Google Video, sitemap per Google Ricerca Codici, sitemap geografiche.

---

<sup>3</sup> XML (eXtensible Markup Language): linguaggio di marcatura estensibile. È un metalinguaggio creato dal W3C con lo scopo di definire la grammatica di diversi linguaggi specifici.

## **2 REALIZZAZIONE DEL SITO WEB “INDICIZZAZIONE E POSIZIONAMENTO NEI MOTORI DI RICERCA”**

In questo terzo capitolo saranno illustrate tutte le fasi che hanno portato alla realizzazione del sito web. Si partirà quindi dall'analisi di blog e siti web per la ricerca del materiale, per passare alla creazione della grafica e del contenuto del sito ed infine ad analizzare gli strumenti messi a disposizione dai motori di ricerca per la segnalazione di un sito web.

### **2.1 ANALISI DI BLOG E SITI WEB**

Il primo passo nella realizzazione del sito web oggetto della tesi è stato l'analisi di un campione di siti web e blog che affrontano il tema dell'indicizzazione e delle diverse tecniche SEO che aiutano a migliorare il posizionamento di un sito web all'interno della SERP.

Il lavoro realizzato trae spunto in particolare dai seguenti siti web:

- [www.giorgiotave.it](http://www.giorgiotave.it): sito internet che affronta il problema del posizionamento di un sito web attraverso diversi aspetti: il sito è infatti costituito, tra le numerose sezioni, da un blog, una directory e un wiki<sup>4</sup>, che si sono rivelati fondamentali per la realizzazione del progetto. Tale sito tratta i diversi aspetti dell'ottimizzazione dei siti web e illustra il funzionamento dei principali motori di ricerca;

---

<sup>4</sup> Sito web i cui contenuti sono sviluppati e modificati in collaborazione da tutti coloro che ne hanno accesso. Scopo fondamentale di questa tipologia di siti web è quello di condividere, scambiare e ottimizzare la conoscenza in modo collaborativo. L'esempio di wiki più famoso è [www.wikipedia.org](http://www.wikipedia.org).

- [www.seoutilility.com](http://www.seoutilility.com): sito web che mette a disposizione dell'utente molti strumenti utili per calcolare diversi parametri di riferimento essenziali nell'attività di web marketing, come il Keywords Density Analyzer<sup>5</sup> e il PageRank Google Check<sup>6</sup>. Questo sito web è stato molto utile, in fase di realizzazione del contenuto delle pagine web, per l'analisi dei cosiddetti meta tag, dei quali fornisce un'accurata descrizione;
- [www.googlerank.com](http://www.googlerank.com): sito che tratta principalmente i metodi di ottimizzazione di una pagina web per il motore di ricerca Google. Affronta il problema della strutturazione dei meta tag della pagina con l'utilizzo del linguaggio HTML per migliorare il posizionamento; alcune sezioni purtroppo non sono ancora complete;
- [www.simonecarletti.com](http://www.simonecarletti.com): blog che tratta i diversi aspetti dell'ottimizzazione e illustra il funzionamento degli algoritmi dei principali motori di ricerca.

## 2.2 CREAZIONE DELLA SCALETTA ARGOMENTATIVA

Dopo questa prima analisi si è proceduto alla creazione di una scaletta argomentativa che fissasse i punti salienti da affrontare nella realizzazione del sito web:

- Home – Introduzione: che cosa s'intende e quali sono le principali tecniche SEO

---

<sup>5</sup> Keywords Density Analyzer: utility che consente di analizzare la densità delle keywords selezionate in relazione ad una pagina web. L'analisi comprende i dettagli della presenza delle key sui meta, i tag principali, densità singole, combinazioni doppie e triple di key.

<sup>6</sup> PageRank Google Check: tools che permette di analizzare il PageRank di una pagina web. Il PageRank è un algoritmo che assegna un peso numerico ad ogni elemento di un collegamento ipertestuale all'interno di un insieme di documenti, con lo scopo di quantificare la sua importanza relativa all'interno della serie.

- Struttura dei dati – in che modo la struttura di una pagina web può influenzare il suo posizionamento all'interno di una SERP
  - HTML – uso dei tag HTML per migliorare l'ottimizzazione della pagina
  - Metadati – cosa s'intende con il termine metadati e quali sono
    - Alt sulle immagini – accorgimenti per migliorare l'utilizzo di questo tag
    - Meta tag – che cosa sono
      - Meta TITLE – uso, scopo, sintassi
      - Meta DESCRIPTION – uso, scopo, sintassi
      - Meta KEYWORDS – uso, scopo, sintassi
      - Meta ROBOTS – uso, scopo, sintassi
      - Meta PRAGMA – uso, scopo, sintassi
      - Meta REVISIT-AFTER – uso, scopo, sintassi
  - Caratteristiche dei dati
    - Keywords density – che cosa significa densità delle parole-chiave e come può essere sfruttata ai fini dell'indicizzazione
    - BackLink – che cosa sono e come utilizzarli per l'ottimizzazione web
  - Motori di ricerca – differenze tra motori di ricerca e directory
    - Google – breve storia del motore di ricerca e funzionamento dell'algoritmo PageRank
    - Yahoo! – breve storia del motore di ricerca e funzionamento di questo ultimo
    - Live Search (già MSN Search) – breve storia del motore di ricerca e suo funzionamento
    - Ask.com (già Ask Jeeves) – breve storia del motore di ricerca e funzionamento dell'algoritmo Edison<sup>7</sup>

---

<sup>7</sup> Algoritmo Edison: nuovo algoritmo utilizzato dal motore di ricerca Ask.com, composto tra tre parti fondamentali: Teoma, che si occupa di determinare l'argomento trattato dalla pagina

- Virgilio – breve storia del motore di ricerca e suo funzionamento
- Arianna – breve storia del motore di ricerca e suo funzionamento

### 2.3 REALIZZAZIONE DELLA GRAFICA

Una volta stabilita la struttura argomentativa del sito web, si è proceduto alla realizzazione di una grafica idonea ai contenuti trattati.

Poiché lo scopo principale della realizzazione di queste pagine web è quello di ottimizzarne la struttura e il contenuto ai fini di una maggiore visibilità, si è preferito utilizzare una veste grafica molto semplice, che non contenesse filmati Flash<sup>8</sup> (anche se ultimamente i motori di ricerca, Google in primis, stanno cercando di migliorare la situazione, è ancora difficoltoso indicizzarne il contenuto).

La grafica è stata realizzata con l'utilizzo del programma Adobe Photoshop<sup>9</sup>, in modo da ottenere gli oggetti grafici previsti nella pagina web; le parti testuali invece, come per esempio i vari menu presenti, sono stati importati all'interno della pagina web come testo poiché, inserendoli come immagini, non sarebbe stata possibile l'indicizzazione di queste parti.

La grafica realizzata è divisa in due parti: la parte sinistra, che occupa circa i 2/3 della schermata del browser, presenta il menu di navigazione e il contenuto

---

web e di valutare i link d'entrata; DirectHit, che utilizza i click dei visitatori per determinare il rank delle pagine; Tagging, che ha lo scopo di rifinire il lavoro effettuato dai due algoritmi precedenti.

<sup>8</sup> Adobe Flash: prodotto della Adobe Systems. Software grafico che permette di creare animazioni vettoriali, principalmente alla scopo di inserirle in pagine web.

<sup>9</sup> Adobe Photoshop: prodotto della Adobe Systems. Software di grafica raster specializzato nell'elaborazione di immagini digitali con la possibilità di lavorare su più "livelli".

della pagina visitata; la parte destra, che occupa 1/3 della pagina, illustra i risultati ottenuti dalla sperimentazione.

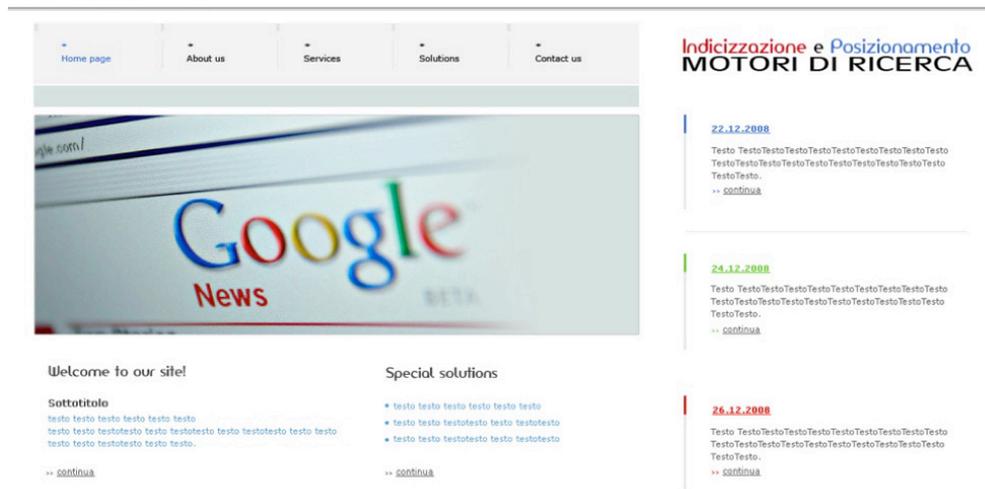


Immagine 2.0: prova grafica

## 2.4 CREAZIONE DEL SITO WEB

Dopo l'implementazione della grafica e la scelta degli argomenti da trattare, si è proceduto alla creazione delle pagine web del sito.

Con l'utilizzo del programma Adobe Dreamweaver<sup>10</sup> si sono quindi create le singole pagine e i relativi contenuti. Completati questi passi, si è proceduto alla creazione dei link sulle voci di menu.

<sup>10</sup> Adobe Dreamweaver: prodotto della Adobe System; programma utilizzato per la realizzazione di pagine web.

## 2.5 CREAZIONE DEI DOMINI E PUBBLICAZIONE IN RETE

Grazie all'aiuto dell'azienda Ariatel Web<sup>11</sup>, nella quale ho portato a termine il tirocinio universitario, si sono creati 6 domini web sui quali è stato duplicato il sito web realizzato.

La fase della scelta dei domini web è di notevole importanza per quanto riguarda l'ottimizzazione, poiché i motori di ricerca danno rilevanza alle parole presenti nell'URL delle pagine; in base a questo fattore quindi si sono scelti indirizzi web significativi, che richiamassero il contenuto delle pagine stesse.

Questi i domini web registrati:

- [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it)
- [www.comesiindicizzaunsitoweb.it](http://www.comesiindicizzaunsitoweb.it)
- [www.lindicizzazione.it](http://www.lindicizzazione.it)
- [www.ilrankingweb.it](http://www.ilrankingweb.it)
- [www.indicizzazioneweb.it](http://www.indicizzazioneweb.it)
- [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it)

Come si può quindi notare da questa lista, si è cercato di utilizzare parole-chiave legate al tema trattato nel sito web, utilizzando spesso termini quali “indicizzazione”, “sito web”, “motori di ricerca”, “algoritmi di ricerca”.

---

<sup>11</sup> Ariatel Web s.r.l.: azienda operante principalmente sul territorio ligure con sede a La Spezia che si occupa della creazione di siti web ad uso commerciale e di web marketing. Sito web: [www.ariatelweb.it](http://www.ariatelweb.it)

Una volta creati gli spazi web, attraverso l'utilizzo del programma FileZilla<sup>12</sup> le pagine web sono state caricate, sui diversi domini, in modo assolutamente omogeneo.

---

<sup>12</sup> FileZilla Client: software gratuito multipiattaforma che permette di trasferire file attraverso il protocollo FTP. Creato un sito FTP, una volta al suo interno è possibile effettuare l'upload o il download dei file.

## **2.6 SEGNALAZIONE DEI SITI WEB AI MOTORI DI RICERCA**

Il passo successivo alla creazione e pubblicazione in rete dei siti web, è la segnalazione di questi ultimi ai motori di ricerca.

Assicurarsi che robot e spider di un motore di ricerca effettuino una o più visite ad un sito web è necessario per incrementare le possibilità che il motore stesso lo indicizzi. L'enorme numero di pagine web presenti in rete può impedire la visita da parte di un crawler o comunque può facilmente incrementare i tempi d'attesa per ottenere buoni risultati per l'ottimizzazione.

La decisione di utilizzare questi strumenti gratuiti è quindi inevitabile ai fini di ottenere risultati utili alla ricerca in tempi brevi.

### **2.6.1 STRUMENTI DI GOOGLE PER I WEBMASTER**

Google mette a disposizione dei suoi utenti una serie di strumenti utili all'attività d'indicizzazione.

Per accedere a tali servizi è necessario prima di tutto creare un account Google; dopodichè, si accede alla cosiddetta bacheca (<https://www.google.com/webmasters/tools>), una pagina in cui è possibile inserire nuovi siti web oppure modificare quelli già inseriti in precedenza.

Una volta inserito l'URL di un nuovo indirizzo web, è possibile accedere ad una pagina di controllo e gestione del sito inserito:

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

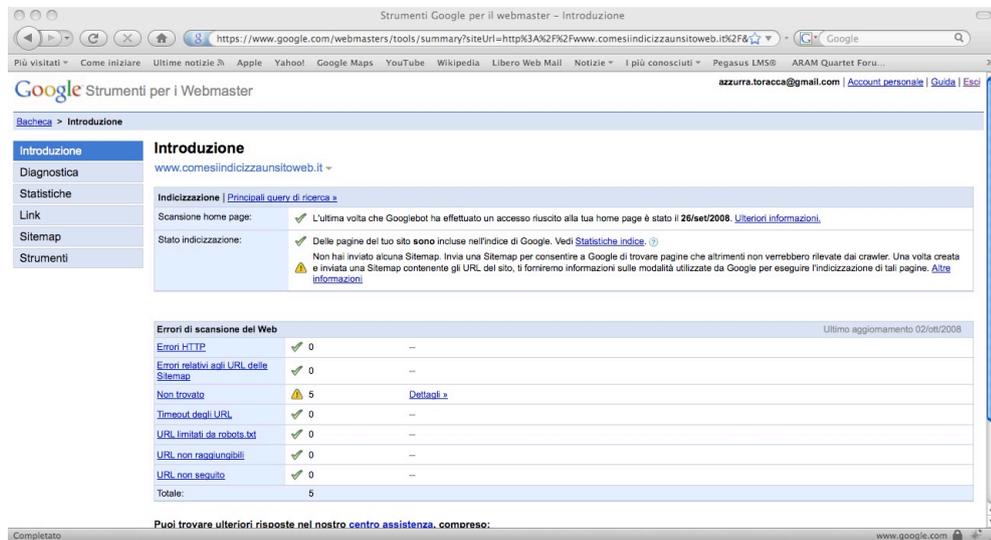


Immagine 2.1: introduzione Strumenti di Google per i Webmaster

La pagina introduttiva mostra una sintesi degli eventuali problemi riscontrati dal crawler di Google nella sua analisi del sito web proposto; il motore di ricerca mette a disposizione dell'utente molti strumenti utili al miglioramento delle proprie pagine ai fini dell'indicizzazione, come la funzione che evidenzia i meta title e i meta description che risultano duplicati all'interno del sito web.

Una volta effettuato l'inserimento del proprio sito web all'interno della bacheca, risulta fondamentale la verifica del sito stesso attraverso la pubblicazione di una pagina HTML vuota (il cui nome è indicato da Google stesso) oppure attraverso l'utilizzo di un meta "verify-v1". Tali strategie occorrono per dimostrare al motore di ricerca che il sito web inserito in bacheca è effettivamente di proprietà di chi l'ha segnalato; la mancata verifica rende vano l'inserimento del sito web in bacheca.

Un'altra funzione fondamentale messa a disposizione dagli strumenti di Google per i webmaster è il caricamento di una o più sitemap, come vedremo in seguito quando parleremo nello specifico della creazione di sitemap per i siti web.

## 2.6.2 YAHOO!: PROPONI L'INCLUSIONE DEL TUO SITO

La proposta d'inclusione di un sito nel motore di ricerca Yahoo! si effettua tramite un link posto nel footer della home page; anche in questo caso è necessario eseguire il login del proprio account o, nel caso in cui non se ne sia in possesso, la registrazione.

Rispetto a Google, Yahoo! permette soltanto l'inserimento dell'URL del proprio feed oppure l'inserimento dell'URL di una pagina web:

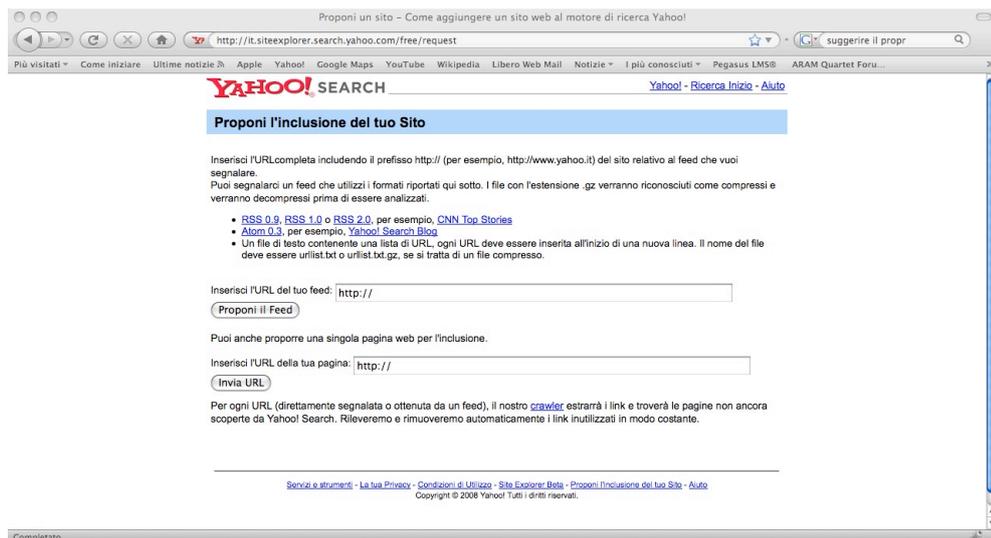


Immagine 2.2: Proponi l'inclusione del tuo sito in Yahoo!

Yahoo! dichiara esplicitamente che non può assicurare né i tempi, né l'effettiva inclusione del sito all'interno del proprio database. Per accedere ad eventuali altri strumenti utili a garantire l'attività di ottimizzazione, in questo crawler occorre effettuare un pagamento mensile.

### 2.6.3 SEGNALA IL TUO SITO A LIVE SEARCH

Per segnalare il proprio sito web a Live Search non è richiesta alcuna registrazione; è sufficiente accedere alla pagina web dedicata alle segnalazioni ed inserire un codice *Captcha*<sup>13</sup> e l'URL della pagina web che vogliamo segnalare.

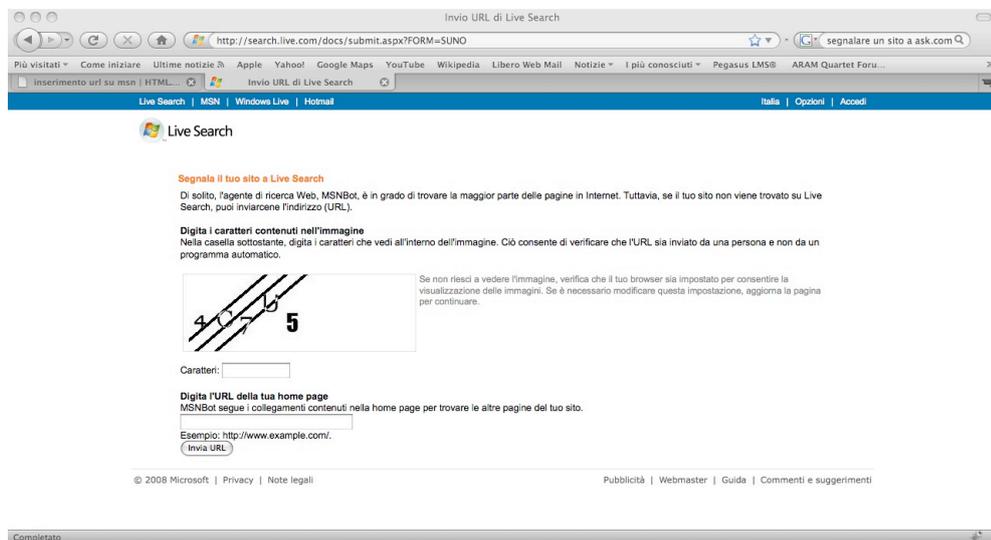


Immagine 2.3: Pagina d'inserimento URL in Live Search

<sup>13</sup> codice Captcha: test che richiede all'utente di inserire una sequenza di lettere o numeri che appaiono distorti od offuscati sullo schermo. Scopo di questo test è quello di verificare che l'utente con cui interagisce il programma sia umano e non un computer.

## 2.6.4 IL MIOASK PER L'INSERIMENTO DI PAGINE WEB

Per segnalare pagine web al motore di ricerca Ask.com è sufficiente prima di tutto accedere alla homepage del motore di ricerca (<http://it.ask.com> per quanto concerne la versione italiana); sulla destra rispetto alla form di ricerca è riportato un menu verticale con diverse voci (web, immagini, conversione valute, ecc...), tra le quali è presente il link MioAsk.

Accedendo a questa pagina viene richiesto il login o, nel caso in cui non si abbia un account su Ask.com, la registrazione gratuita al motore di ricerca.

Effettuato il login, si accede al pannello di MioAsk, nel quale è possibile avere sotto controllo le pagine web inserite in precedenza:

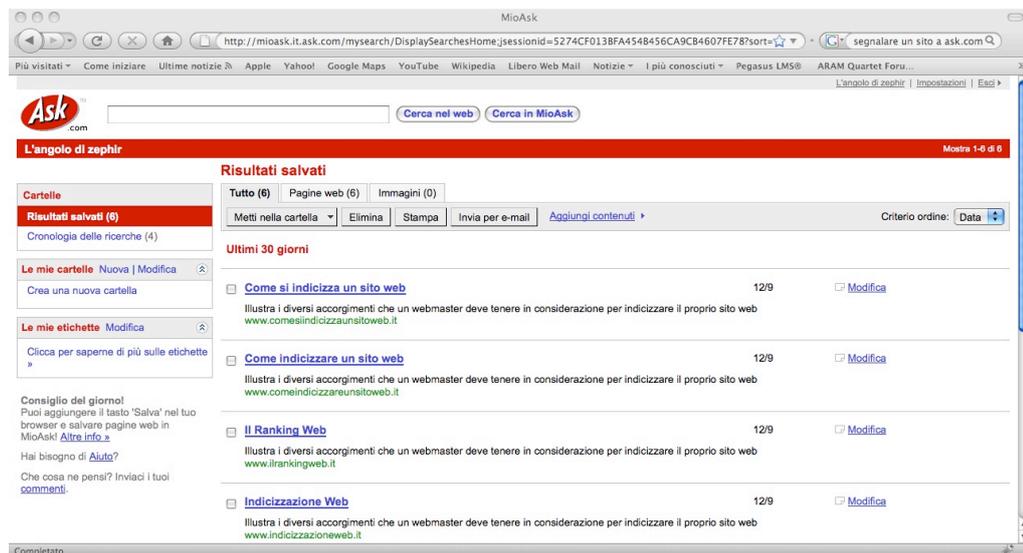


Immagine 2.4: pannello di controllo del MioAsk

Per inserire un nuovo URL, è sufficiente inserire titolo del sito web, URL e una breve descrizione del contenuto del sito stesso:

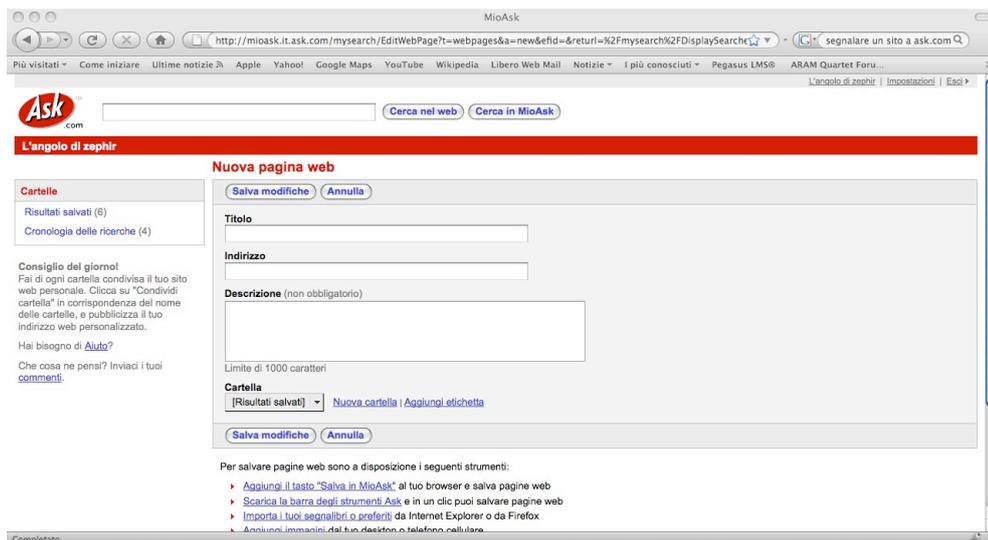


Immagine 2.5: schermata d'inserimento URL nel MioAsk

In questa pagina, per l'inserimento di siti web, è richiesta la digitazione del titolo del sito, dell'URL e di una breve descrizione del contenuto del sito stesso.

## 2.6.5 MOTORI DI RICERCA VIRGILIO E ARIANNA

Questi due motori di ricerca italiani, originariamente nati come directory autonome, oggi sfruttano, totalmente (Arianna) o principalmente (Virgilio), i risultati forniti dal motore di ricerca Google. Nel caso di questi motori di ricerca, non è stato quindi necessario utilizzare strumenti dedicati per segnalare un sito web.

### **3 PREPARAZIONE ED ESECUZIONE DEGLI ESPERIMENTI**

All'interno di questa sezione saranno illustrati tutti i passi compiuti al fine di ottimizzare i siti web per l'indicizzazione da parte dei motori di ricerca. Si parlerà prima di tutto della scelta delle parole chiave, dei metodi di ottimizzazione scelti ed infine di come i diversi domini sono stati modificati in base alle tecniche prescelte.

#### **3.1 SCELTA DELLE PAROLE CHIAVE**

Ottimizzare un sito web richiede prima di tutto una fase di analisi dei contenuti per determinare con criterio una lista di parole con cui si desidera acquisire posizioni all'interno di una SERP.

Risulta quindi una fase fondamentale per ottenere buoni risultati dall'applicazione di una o più tecniche SEO; applicare con precisione tecniche di ottimizzazione ad un sito web per parole chiave sbagliate può infatti non portare a nessun risultato. È quindi fondamentale analizzare attentamente i contenuti delle pagine e assegnare ad ognuna di esse parole chiave che compaiono effettivamente all'interno della pagina o che comunque siano coerenti con gli argomenti in essa trattati.

Si è quindi eseguita un'accurata analisi dei contenuti delle pagine ed è stata stilata una lista di circa 30 parole chiave, utilizzata per l'ottimizzazione di tutti i domini creati:

<b>PAROLE CHIAVE</b>		
Ask.com	Motori di ricerca	Algoritmi di ricerca
Linguaggio HTML	Alt sulle immagini	Motori e algoritmi di ricerca
Arianna	BackLink	Posizionamento siti web
Google	Link popularity	Metodi di indicizzazione
PageRank	Tecniche SEO	Ottimizzazione siti web
Parole chiave	Keywords density	Tag HTML
Meta tag	metadati	Meta title
Meta pragma	Meta keywords	Meta description
Meta robots	Meta revisit-after	Directory
Algoritmo Edison	yahoo	virgilio
Algoritmo Teoma <sup>14</sup>	ExpertRank <sup>15</sup>	

### 3.2 METODI DI OTTIMIZZAZIONE UTILIZZATI

Poiché il problema di come migliorare la visibilità di un sito web è un tema molto attuale e sul quale sono tuttora aperti dibattiti, esistono pareri più o meno contrastanti sull'effettiva efficacia delle tecniche SEO utilizzate dai webmaster. Dopo un'accurata analisi, è stato deciso di prendere in considerazione gli stru-

---

<sup>14</sup> Algoritmo Teoma: vecchio algoritmo di ricerca utilizzato da Ask.com, attualmente integrato nell'Algoritmo Edison (v. nota 5). Il suo scopo è quello di individuare l'argomento principale trattato in una pagina web e di verificare i link in entrata.

<sup>15</sup> ExpertRank: algoritmo di ricerca storico utilizzato dal crawler Ask.com, attraverso il quale vengono assegnati diversi valori per ogni pagina web indicizzata.

menti maggiormente utilizzati a questi scopi, benché su alcuni di essi, come per esempio i meta tag, molti esperti del settore abbiano manifestato pareri contrastanti riguardo la loro effettiva utilità.

La scelta di tecniche su cui ancora oggi si discute è stata dettata dalla volontà di sperimentarle direttamente e capire quali possono essere ritenute realmente efficaci.

I metodi di miglioramento della visibilità presi in considerazione sono quindi i seguenti:

1. Meta tag
2. Link e BackLink
3. Linguaggio HTML
4. Sitemap

Questa la suddivisione delle tecniche SEO sui domini precedentemente creati:

- [www.comesiindicizzaunsitoweb.it](http://www.comesiindicizzaunsitoweb.it): sito web nel quale sono stati inseriti una serie di meta tag;
- [www.ilrankingweb.it](http://www.ilrankingweb.it): dominio in cui sono stati utilizzati come ottimizzatori i link in uscita e in entrata (BackLink) dal sito;
- [www.indicizzazioneweb.it](http://www.indicizzazioneweb.it): dominio in cui è stata inserita una sitemap, linkata da ogni pagina del sito;
- [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it): sito web in cui sono stati utilizzati i tag HTML come metodo di indicizzazione;
- [www.lindicizzazione.it](http://www.lindicizzazione.it): dominio non modificato da nessuna tecnica;
- [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it): sito a cui sono stati applicati tutti i metodi di indicizzazione visti nei precedenti domini

### **3.3 MODIFICA DELLA STRUTTURA DEI DOMINI AI FINI DELL'INDICIZZAZIONE**

#### **3.3.1 WWW.COMESIINDICIZZAUNSIOWEB.IT**

##### *Uso dei meta tag ai fini dell'ottimizzazione*

L'ottimizzazione di questo dominio web è stata effettuata attraverso l'inserimento, all'interno di <head></head>, di una serie di meta tag, gli stessi illustrati all'interno del capitolo 2: meta title, meta description, meta keywords, meta robots, meta pragma, meta revisit-after.

Nel meta title sono state inserite, per ogni pagina, alcune parole chiave dell'elenco deciso in precedenza.

Come già spiegato brevemente nell'introduzione a questa relazione, è molto importante scegliere con criterio le parole chiave da inserire in ogni pagina web: sarà logico quindi inserire all'interno di questo meta parole che compaiono nel contenuto della pagina e che indicano l'argomento principale affrontato al suo interno. Nella pagina dedicata ai motori di ricerca, ad esempio, sono state inserite parole chiave come *motori di ricerca*, *algoritmi di ricerca*, *motori e algoritmi di ricerca* e, all'interno di ogni singola pagina dedicata ad uno specifico crawler o directory, sono stati richiamati i nomi di questi.

Anche il contenuto del meta tag description è determinato in base al medesimo criterio: cercare di dare una spiegazione all'utente di che cosa troverà visitando la pagina web, tentando in ogni caso di riutilizzare le parole chiave concordate in precedenza.

La spiegazione deve essere semplice e concisa e rappresentare il contenuto effettivo della pagina, cercando di inserire le parole salienti nei primi posti del content del meta.

Il meta keywords ha lo scopo di contenere ed indicare tutte le parole-chiave in base alle quali si vorrebbe indicizzare una pagina web.

Anche in questo caso è importante inserire ai primi posti le parole-chiave più importanti per quella pagina. Per cercare di rimarcare ulteriormente le frasi per l'ottimizzazione, è possibile anche ripetere le singole parole separate: ovvero, nel caso di parole-chiave complesse, come *algoritmo Edison*, è utile inserire come parole chiave anche *algoritmo* e *Edison*.

Questo soprattutto perché, quando si tratta di frasi (in particolar modo quando le parole sono più di due), è possibile che l'utente non digiti esattamente la frase così com'è stata formulata dal webmaster per l'ottimizzazione. In questo modo è quindi possibile aumentare i casi in cui il nostro sito o la nostra pagina web rientri all'interno dei risultati di un motore di ricerca.

Poiché il meta robots ha lo scopo di indicare il comportamento che gli spider di un motore di ricerca devono tenere nel caso in cui visitino un sito web, è bene indicare come contenuto di questo tag "*index, follow*". In questo modo infatti s'indica allo spider di indicizzare i contenuti della pagina (*index*) e di seguire tutti i link presenti all'interno della pagina stessa (*follow*); queste impostazioni possono migliorare l'indicizzazione da parte di crawler e directory.

Infine è stato inserito il meta revisit-after per indicare allo spider dopo quanti giorni si desidera ricevere una sua nuova visita. Indicare un numero di giorni troppo bassi può essere controproducente, soprattutto se in successive visite lo spider continua a non registrare modifiche al contenuto della pagina. Allo stesso modo, indicare un numero troppo alto, può impedire la tempestiva registrazione delle modifiche riportate alle pagine da parte del crawler.

All'interno del meta revisit-after del sito web in questione, il tempo d'attesa è stato impostato a 10 giorni.

Il meta pragma non è stato inserito come meta attivo nelle pagine web del sito in quanto, nell'accezione analizzata, il suo scopo è quello di evitare il salvatag-

gio dei contenuti della pagina nella memoria cache del computer. Poiché è consigliabile scegliere questa opzione soprattutto nei casi di siti in continuo aggiornamento, si è deciso di ometterlo nella nostra lista.

### 3.3.2 WWW.ILRANKINGWEB.IT

#### *Usa dei link d'entrata e d'uscita da una pagina web*

Questo dominio web è stato ottimizzato per l'indicizzazione attraverso l'utilizzo di link interni, veri e propri rimandi alle altre pagine interne del sito stesso e talvolta anche a siti esterni, utilizzando sempre le parole-chiave come ancora per i medesimi link.

Ma, con indicizzazione attraverso link, s'intendono anche i cosiddetti BackLink, ovvero i link che un sito web riceve da siti esterni. Guadagnare BackLink, soprattutto guadagnare buoni BackLink, necessita di molto tempo, ma una pratica utile è quella di segnalare ed inserire i nostri siti all'interno di directory.

Il mondo delle directory è ormai vastissimo, n'esistono di molteplici tipi: a pagamento e non, che richiedono un link di ritorno obbligatorio, facoltativo oppure nessun link. È stato quindi opportuno determinare una serie di criteri per effettuare una scelta di directory in cui inserire il nostro sito web.

Ricercando su Google "directory italiane", il motore di ricerca restituisce una serie di pagine web contenenti una lista di directory.

Visitando alcune di queste liste si sono potute conoscere una serie di directory, tra le più conosciute e utilizzate del web. Naturalmente, molte di queste richiedevano il pagamento di una somma per la sottomissione del proprio sito; queste sono state quindi scartate. Si è inoltre cercato di evitare anche quelle directory che richiedevano l'inserimento di un link di ritorno; questo soprattutto perché l'inserimento di un numero eccessivo di link, in particolare diretti a siti web che non trattano un argomento simile a quello del nostro sito, possono determinare, da parte di Google in particolare, un posizionamento peggiore all'interno di una SERP. Questo è dettato dal fatto che in passato, per ottenere

buone posizioni, i webmaster inserivano nei propri siti una grande quantità di link, spesso del tutto estranei all'argomento del sito che li linkava. Ad oggi quindi i motori di ricerca non considerano soltanto la quantità di link presenti in una pagina, ma anche la loro qualità, cioè la coerenza e la popolarità dei siti linkati.

Presi in considerazione questi fattori, sono state scelte sette directory in cui inserire il nostro sito web.

#### *Giorgio Tave Directory*

La scelta di inserire il sito web in questa directory è stata dettata dal fatto che, durante la realizzazione del progetto, questo sito web è stato frequentemente consultato per ricavarne informazioni sulle tecniche di ottimizzazione.

La registrazione in questa directory richiede l'inserimento obbligatorio del titolo del sito internet, l'URL del sito stesso, una breve descrizione di che cosa tratta il sito internet, la categoria in cui si vuole inserire il sito web (con eventuali sottocategorie) e un'e-mail di riferimento. È inoltre possibile inserire una serie di campi supplementari.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

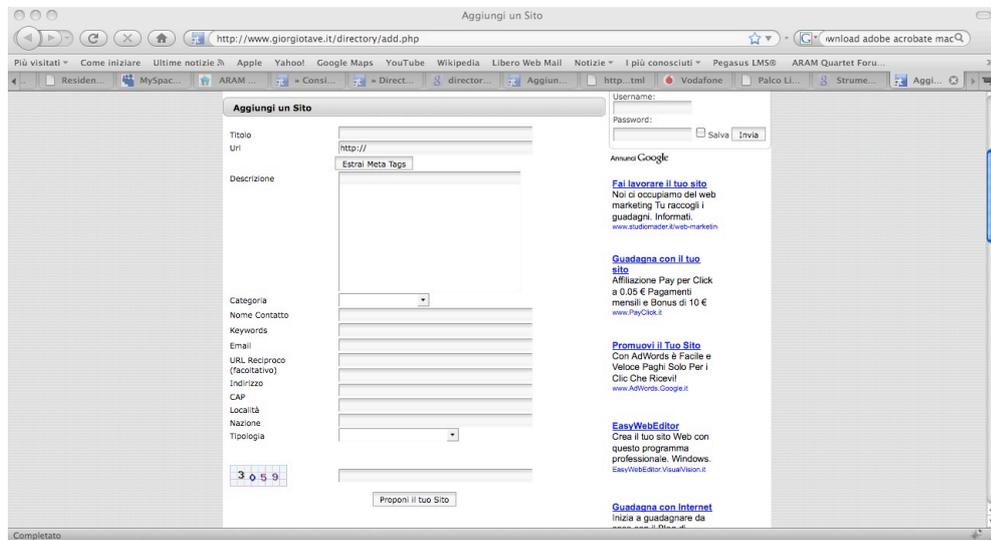


Immagine 3.0: Giorgio Tave directory

La categoria in cui è stato inserito il sito web è “Internet – Posizionamento Motori di Ricerca – Guide”.

### *SegnalaSito Directory*

SegnalaSito è una delle directory italiane più visitate, presente nelle diverse liste consultate; questa infatti ha un valore di PageRank uguale a 5, molto alto se si considera che Dmoz, una delle directory più famose, ha un PageRank pari a 7.

Si tratta di una web directory gratuita divisa in 11 categorie, con relative sottocategorie tematiche, che permette di eseguire una ricerca all'interno dei siti registrati nel suo archivio.

Anche in questa directory, l'inserimento di un sito web richiede la specifica di alcune informazioni: titolo della pagina, URL, categoria d'appartenenza (ed

eventuale sottocategoria), breve descrizione del contenuto del sito, proprio nome e indirizzo mail.

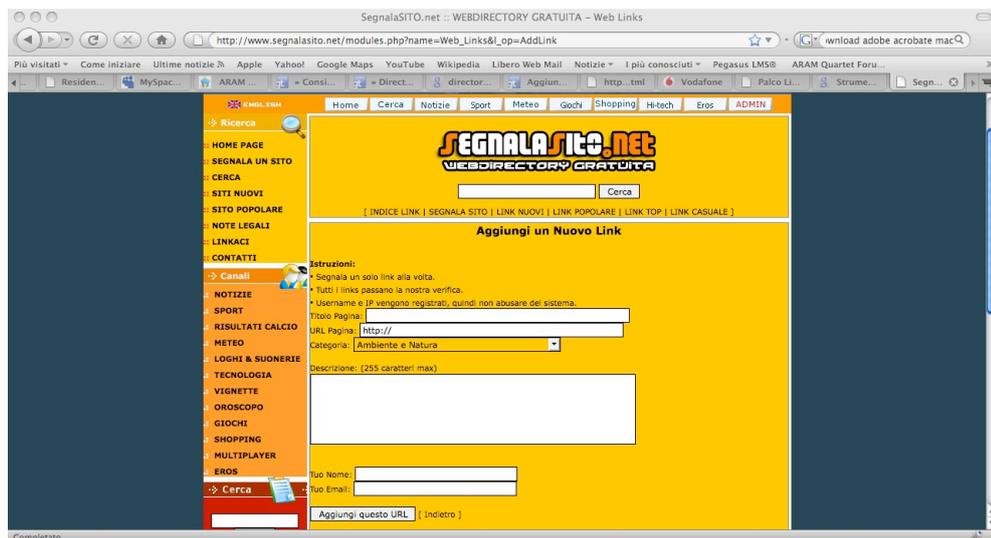


Immagine 3.1: inserimento di un sito nella directory SegnalaSito.net

### *Dmoz Directory*

Dmoz è una directory multilingue divisa in 15 categorie con eventuali sottocategorie per ognuna di esse.

Directory presente sul web ormai da alcuni anni e ritenuta molto valida da coloro che si occupano di migliorare il posizionamento dei propri siti web all'interno dei motori di ricerca.

Il lavoro di questa directory è totalmente gestito da persone che controllano i siti proposti dagli utenti per la sottomissione e che decidono se inserirli o meno all'interno del proprio archivio. L'inserimento di un sito web all'interno di Dmoz può quindi richiedere un tempo d'attesa che va da alcune settimane ad un periodo indefinito: questo tipo d'approccio, tuttavia, permette un accurato esame di ogni singola pagina inserita all'interno di Dmoz.

# Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

Proponi un sito a Open Directory

http://www.dmoz.org/cgi-bin/add.cgi?where=World/Italiano/Computer/Internet

Più visitati | Come iniziare | Ultime notizie | Apple | Yahoo! | Google Maps | YouTube | Wikipedia | Libero Web Mail | Notizie | I più conosciuti | Pegasus LMSB | ARAM Quartet Foru... | MySpac... | ARAM... | Cons... | Direct... | director... | Aggiun... | http...tml | Vodafone | Palco Li... | Strume... | General... | Prop...

### Proponi un sito a Open Directory

Grazie per l'interesse dimostrato in Open Directory. Proporre un sito è facile, ma prima di procedere con il modulo di invio, ti chiediamo di fare due cose.

1. Per favore prendi un momento per dare un sguardo alle nostre [regole riguardo alle proposte di siti](#). È importante che tu capisca queste regole. La mancata comprensione e rispetto di queste regole generalmente comporterà il rifiuto della proposta di inserimento.
2. Controlla per essere sicuro che questa sia la sola categoria nella quale pensi il tuo sito dovrebbe essere elencato. Open Directory ha un ricco albero di temi, e sarà di aiuto per tutti se cerchi la migliore categoria. Questo aiuta a rendere veloce la nostra recensione del tuo sito.

**Esempio:**

Un sito sulla Chiesa Cattolica dovrebbe essere proposto in:

**Top: World: Italiano: Società: Religione e Spiritualità: Cristianesimo: Denominazioni: Chiesa Cattolica e non**

**Top: World: Italiano: Società: Religione e Spiritualità: Cristianesimo**

Questa è una distinzione importante nel mondo della categorizzazione web, e assicura un veloce esame del tuo sito.

Per favore fai attenzione: Noi non siamo un motore di ricerca e siamo orgogliosi di essere altamente selettivi. Noi non accettiamo tutti i siti, quindi, per favore non prenderla personalmente se il tuo sito non sarà accettato.

**Categoria:** **World: Italiano: Computer: Internet**

I siti in lingua diversa dall'italiano devono essere sottoposti nella categoria linguistica appropriata di [World](#).

Può essere utile prendere visione delle [Regole generali per l'inserimento dei siti](#) che valgono per tutte le categorie della struttura di [World/Italiano](#).

⚠ Sei sicuro che questa categoria sia quella che meglio corrisponde al tuo sito? Nel dubbio, ricerca nella directory la categoria più appropriata.

---

**URL del Sito:**

Che tipo di link è questo?  Regolare  PDF  RSS  Atom

⚠ URL significa Uniform Resource Locator, ovvero l'indirizzo del tuo sito. Esempio: <http://dmoz.org>

- Non proporre siti mirror.
- Non proporre una URL che contiene solo materiale uguale o simile a quello di altri tuoi siti che sono già elencati nella directory. Richieste di inserimento multiple dello stesso sito o di siti correlati possono comportare l'esclusione o la cancellazione di tutti i siti affiliati.
- Non mascherare il sito che proponi inserendo lo stesso URL più di una volta. Esempio: <http://www.dmoz.org> e <http://www.dmoz.org/index.html>.
- Non proporre siti che reindirizzano ad un altro sito.
- Open Directory ha una politica contraria all'inclusione di siti con contenuto illegale. Esempi di materiale illegale includono pornografia infantile, diffamazione, materiale che viola qualche diritto di proprietà intellettuale, o materiale che specificamente sostiene, sollecita o favorisce attività illegali (come frodi o violenze).
- Non proporre siti "in costruzione".
- Proponi siti pornografici nella categoria appropriata sotto Adult.
- Proponi siti in lingua diversa dall'inglese nella categoria appropriata sotto World.
- Non proporre siti costituiti prevalentemente da affiliate link.

---

**Titolo del Sito:**

⚠ Dai al tuo sito un titolo breve e descrittivo.

- Opta sempre per il nome ufficiale del sito.
- Non usare il TUTTO MAIUSCOLO.
- Evita il linguaggio promozionale nel titolo.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

Descrizione del Sito:

▲Rendi una breve descrizione del tuo sito: non più lunga di 25-30 parole. Una descrizione obiettiva e ben scritta renderà più facile l'inserimento del vostro sito.

- Non usare alcuna tag HTML.
- Scrivi frasi complete e grammaticalmente corrette, con la giusta punteggiatura.
- Non scrivere TUTTO MAIUSCOLO nella descrizione.
- Evita di usare un linguaggio promozionale o stringhe di keyword e chiavi di ricerca. Parole e frasi come "bellissimo" o "il più bel sito" verranno rimosse.

Il Tuo Indirizzo

E-mail:

User Verification: ▲Please verify the text in the image. The text in the image and the text you submit in the input field must match.

Can't read this text? [Get New Text](#)

Nmp5Mf3M

Completato

Immagine 3.2: sottomissione sito web alla directory Dmoz

### *Mg Quadro Directory*

Directory in lingua italiana che suddivide le pagine contenute nel proprio archivio in 14 categorie, suddivise a loro volta in sottocategorie.

Questa directory vanta un funzionamento secondo due caratteristiche fondamentali:

1. Link diretto: “il titolo proposto per la risorsa diventa l’anchor diretto del link della risorsa stessa, senza la presenza di script di reindirizzamento. Questo consente non solo un passaggio di PageRank, ma anche la possibilità di beneficiare semanticamente dei termini presenti nel titolo”;
2. Page navigation: “tutte le sottocategorie consistono di un’unica pagina senza la navigazione in pagine successive che disperdono il PageRank”.

La sottomissione di un sito web a questa directory richiede l’inserimento di titolo, descrizione, URL.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

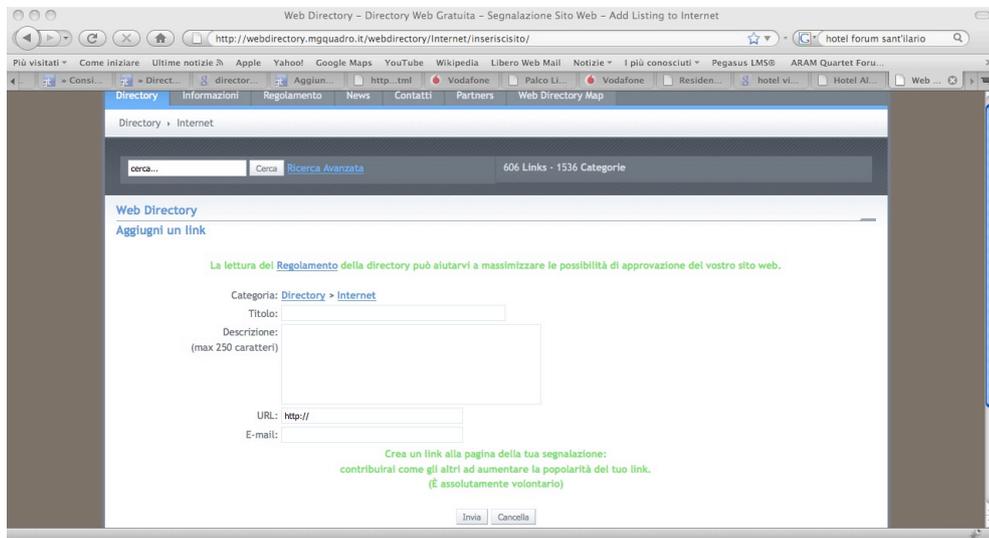


Immagine 3.3: sottomissione di un sito web alla directory

### *SeoTag Directory*

Sito web interamente dedicato al web marketing che, oltre alla directory, contiene guide e articoli sull'argomento.

La directory è suddivisa in 16 categorie; il sito web è stato inserito in "Internet – Guide Web Marketing". Le informazioni richieste per la sottomissione sono titolo, descrizione e URL.

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

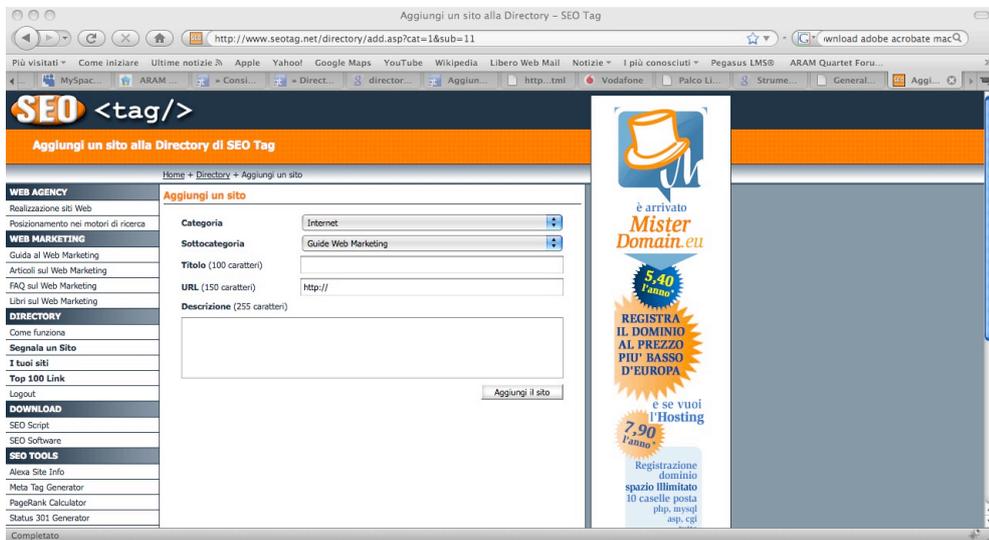


Immagine 3.4: inserimento di un sito web all'interno della directory SeoTag

*Infiniti Spazi Directory*

Directory italiana che permette di inserire i propri siti web gratuitamente previo inserimento di un link reciproco.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

The image shows a web browser window displaying the 'Inferispazi Directory - Inserisci link' page. The browser's address bar shows the URL 'http://directory.infinitispazi.com/submit.php'. The page features a navigation menu with links like 'Inserisci link', 'Inserisci Articolo', and 'Ultimi Link'. On the left, there are sections for 'CATEGORIES' and 'STATISTICS'. The main content area is a form for submitting a link, with fields for 'Titolo', 'URL', 'Descrizione', 'META Keywords', 'META Description', 'Il Tuo Nome', 'La Tua Email', 'Categoria', and 'URL del Link Reciproco'. A CAPTCHA image is visible at the bottom of the form. The page footer contains copyright information for Infinitispazi Directory © 2008.

Immagine 3.5: inserimento sito web nella directory Infiniti Spazi

### *My Network Directory*

Directory in lingua italiana che richiede una registrazione gratuita prima dell'inserimento dei propri siti. Una volta effettuato il login, si accede al pro-

prio pannello di controllo personale, in cui è possibile modificare i siti web già inseriti oppure aggiungerne di nuovi.

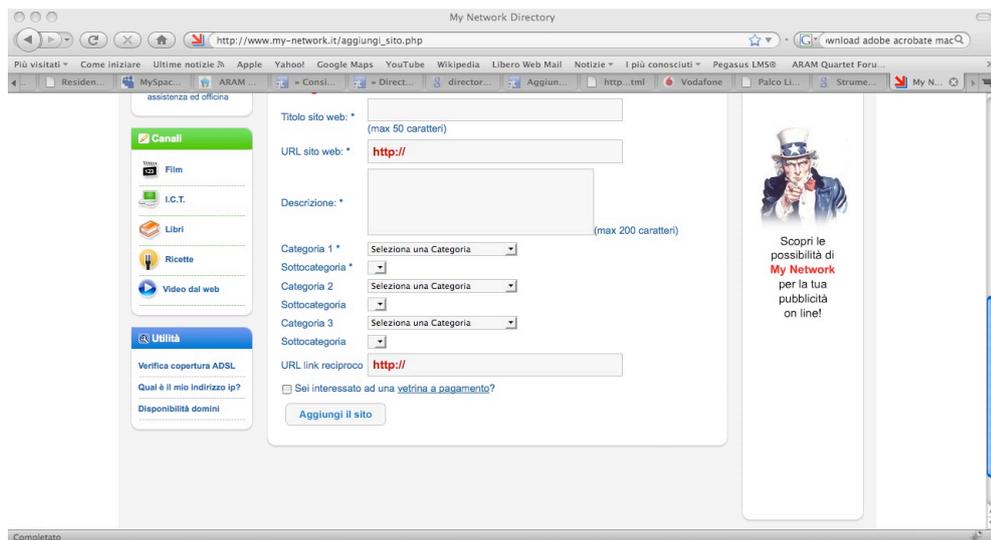


Immagine 3.6: inserimento sito web all'interno della directory My network

Per l'inserimento di un sito web all'interno della directory sono richiesti:

- titolo
- URL
- descrizione
- categoria e sottocategoria 1: Informatica Internet e Telefonia – Risorse
- categoria e sottocategoria 2: Informatica Internet e Telefonia – Web Marketing

È inoltre prevista la possibilità di inserire, facoltativamente, un URL reciproco.

### 3.3.3 WWW.INDICIZZAZIONEWEB.IT *Creazione di una Sitemap*

Come illustrato nel capitolo 1, esistono essenzialmente due tipi di sitemap: testuale e in formato XML; all'interno del sito web sono state inserite entrambe le tipologie.

La sitemap in formato testuale è stata realizzata attraverso l'uso di un elenco puntato che evidenzia le gerarchie esistenti tra le pagine e su cui sono stati inseriti i link alle pagine corrispondenti per ogni voce della lista:

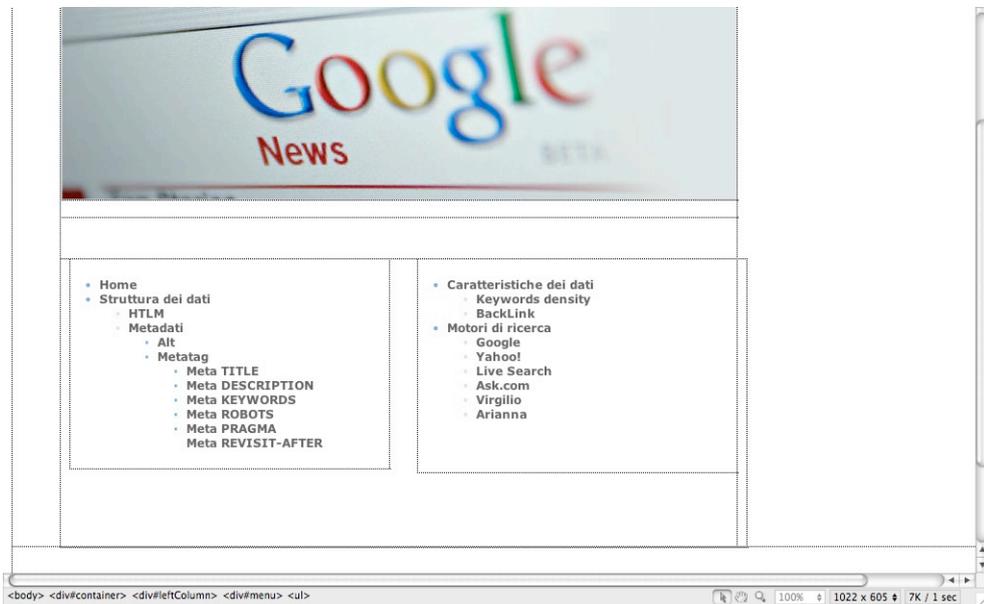


Immagine 3.7: sitemap del sito web

Per quanto riguarda invece la creazione della sitemap in formato XML, è stato utilizzato un programma apposito, SOFT Plus GSite Crawler<sup>16</sup>, completamente gratuito.

Inserito l'URL del sito web per il quale si desidera generare la mappa, saranno forniti una serie di file in formato XML e in formato robots che dovranno essere pubblicati in rete all'interno del sito web.

Una volta generati e pubblicati i file, è importante indicare l'esistenza del file XML a Google; è sufficiente quindi accedere alla pagina degli Strumenti per i Webmaster e, nella sezione apposita, inserire il nome del file generato.

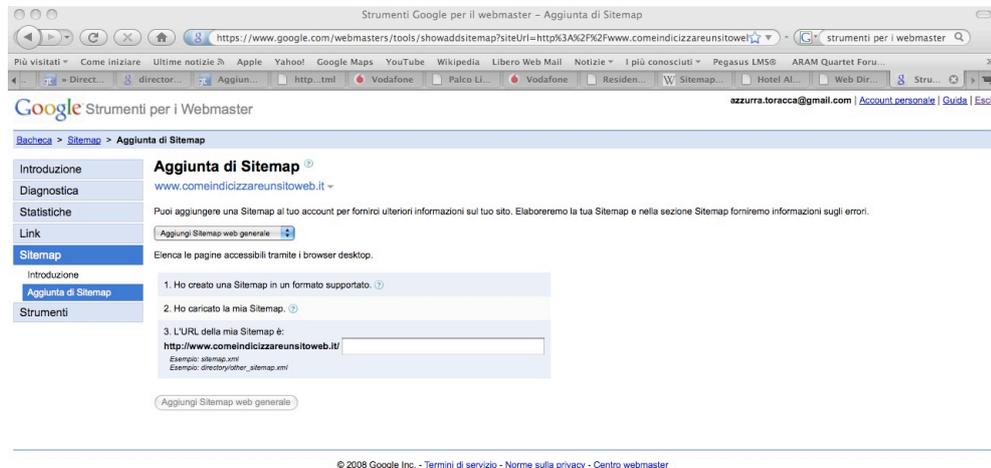


Immagine 3.8: aggiungere una sitemap in formato XML agli Strumenti di Google per i Webmaster

<sup>16</sup> Soft Plus GSite Crawler: prodotto della SOFTplus Entwicklungen GmbH. Programma che permette la creazione di sitemap esportabili in diversi formati.

### 3.3.4 WWW.MOTORIEALGORITMIDIRICERCA.IT

#### *Uso dei tag HTML per l'ottimizzazione*

Come illustrato nel primo capitolo, utilizzare un codice HTML pulito e chiaro può aiutare il posizionamento di un sito web all'interno di una SERP. Inoltre, inserire le parole chiave con cui si vuole indicizzare una pagina web in determinati tag HTML permette di dar loro maggiore importanza rispetto ad altre.

In questa sperimentazione sono stati quindi utilizzati in modo strategico alcuni tag messi a disposizione del linguaggio HTML. In particolare, in ogni pagina web del sito è stato inserito un tag `<h1></h1>` che contiene le parole chiave per cui si vuole indicizzarla. Poiché ogni pagina era già stata provvista di un determinato titolo, si è proceduto in modo che il contenuto di questo tag, una volta caricata la pagina da un qualunque browser, non fosse visibile all'utente. All'interno del contenuto della pagina, si è invece proceduto all'utilizzo dei tag `<b></b>` e `<em></em>` per evidenziare le parole chiave con cui volevamo che il nostro sito internet fosse indicizzato. La scelta di questi tag è stata dettata dal fatto che i motori di ricerca danno molta importanza alle parole evidenziate in grassetto o in corsivo.

Per rimarcare il contenuto di alcune frasi, è stato inoltre utilizzato il tag `<h3></h3>`, con classe "ind", attraverso la quale sono state date caratteristiche testuali del tutto identiche al contenuto dei tag `<p></p>`, presenti nel corpo della pagina stessa.

In questo modo quindi le parole-chiave presenti all'interno di questo tag rivestono un'importanza maggiore rispetto alle altre, nonostante il loro aspetto, nel momento in cui la pagina è visualizzata da un browser web, risulti identico a tutte le altre.

### 3.3.5 WWW.LINDICIZZAZIONE.IT

*Sito web non modificato da alcun metodo di ottimizzazione*

Questo dominio web è stato volutamente lasciato invariato, senza l'aggiunta di nessuna tecnica di ottimizzazione, al fine di confrontarlo con il lavoro finale effettuato su tutti quei siti che sono stati invece sottoposti a modifiche. Scopo di questa scelta è quello di mostrare come un sito, costruito senza nessun accorgimento nei confronti dei motori di ricerca e delle loro tecniche di catalogazione di un sito web, difficilmente può raggiungere una discreta posizione all'interno di una potenziale SERP di uno più motori di ricerca.

### 3.3.6 WWW.COMEINDICIZZAREUNSIOWEB.IT

*Sito web ottimizzato attraverso l'uso delle tecniche SEO inserite separatamente negli altri siti web*

Il sesto ed ultimo dominio è stato modificato con l'aggiunta di tutte le tecniche di ottimizzazione applicate singolarmente ai precedenti siti web.

Il mondo del web marketing è ancora molto incerto e, come dimostrano i frequenti dibattiti sollevati sull'argomento, non esistono tecniche univoche che possono assicurare la buona indicizzazione di una pagina web. É quindi difficile che l'applicazione di una sola tecnica SEO possa garantire il buon posizionamento di un sito.

Scopo della creazione di questo sesto dominio è quindi quello di dimostrare che, applicando con giudizio diverse tecniche ad uno stesso documento, questo incrementerà maggiormente il suo posizionamento rispetto ad uno modificato tramite l'utilizzo di una sola tecnica.

Questo dominio web è stato quindi modificato seguendo i medesimi procedimenti utilizzati per i quattro siti web illustrati in precedenza: inserimento di

meta tag nell'head della pagina come per [www.comesiindicizzaunsitoweb.it](http://www.comesiindicizzaunsitoweb.it);  
sottomissione ad una serie di directory gratuite come per [www.ilrankingweb.it](http://www.ilrankingweb.it);  
modifica e aggiunta di tag HTML come in [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it);  
inserimento di due tipi di sitemap, testuale e XML, come per [www.indicizzazioneweb.it](http://www.indicizzazioneweb.it).

## **4 RACCOLTA E ANALISI DEI DATI CONCLUSIVI**

In questo capitolo si procederà alla raccolta di una serie di dati statistici: le visite di spider e robot ai domini web creati, il numero di accessi ricevuti dagli utenti e, nel caso di accessi tramite motori di ricerca, le parole chiave che hanno indirizzato le visite verso i siti web.

Saranno inoltre analizzati i dati relativi agli incrementi di posizioni dei siti web per una serie di parole chiave.

Trascorse tre settimane dalla messa online delle modifiche apportate ai domini web ai fini di un'ottimizzazione dei contenuti, sono stati raccolti i dati utili a stilare una serie di conclusioni sul lavoro svolto.

I dati utilizzati a questo scopo sono stati essenzialmente di due tipi: dati statistici su accessi e tipi di accessi da parte di utenti e motori di ricerca; scansioni effettuate al fine di ottenere una lista di parole chiave su cui i domini web sono stati indicizzati e la loro posizione all'interno della lista di risultati dei singoli motori presi in esame.

### **4.1 INFORMAZIONI STATISTICHE RICAVATE DA STAT.ARIATEL.IT**

Grazie all'azienda Ariatel Web s.r.l., che ha permesso la pubblicazione in rete dei siti web realizzati, è stato possibile accedere a una pagina contenente una serie di dati statistici riguardanti visite e accessi.

Per accedere a questi dati, è necessario essere a conoscenza di username e password del profilo (appositamente creato dai tecniche Ariatel) e l'URL del dominio web. Ovviamente, i primi due dati richiesti risultano invariati per tutti i siti web mentre nella form "dominio" è richiesto l'inserimento dell'indirizzo web di ogni singolo sito internet.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

Eseguito l'accesso alle statistiche, la pagina web si presenta come segue:

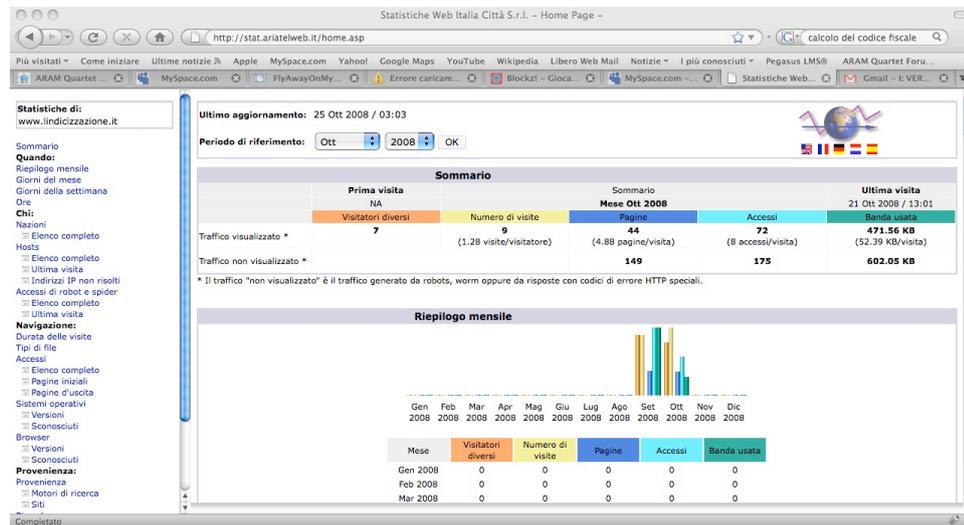


Immagine 4.0: home page di stat.ariatelweb.it

Attraverso questa pagina, è possibile conoscere, ad esempio, il numero di accessi effettuati verso quel determinato indirizzo web, in quali giorni sono stati effettuati tali accessi e, di conseguenza, i giorni della settimana in cui si hanno più accessi. Altro dato che può rivelarsi utile, soprattutto nel campo del web marketing e delle campagne d'indicizzazione, è quello riguardante le nazioni da cui sono stati effettuati gli accessi.

Una serie di dati statistici fondamentali per la ricerca effettuata in quest'ambito riguardano gli accessi di robot e spider:

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

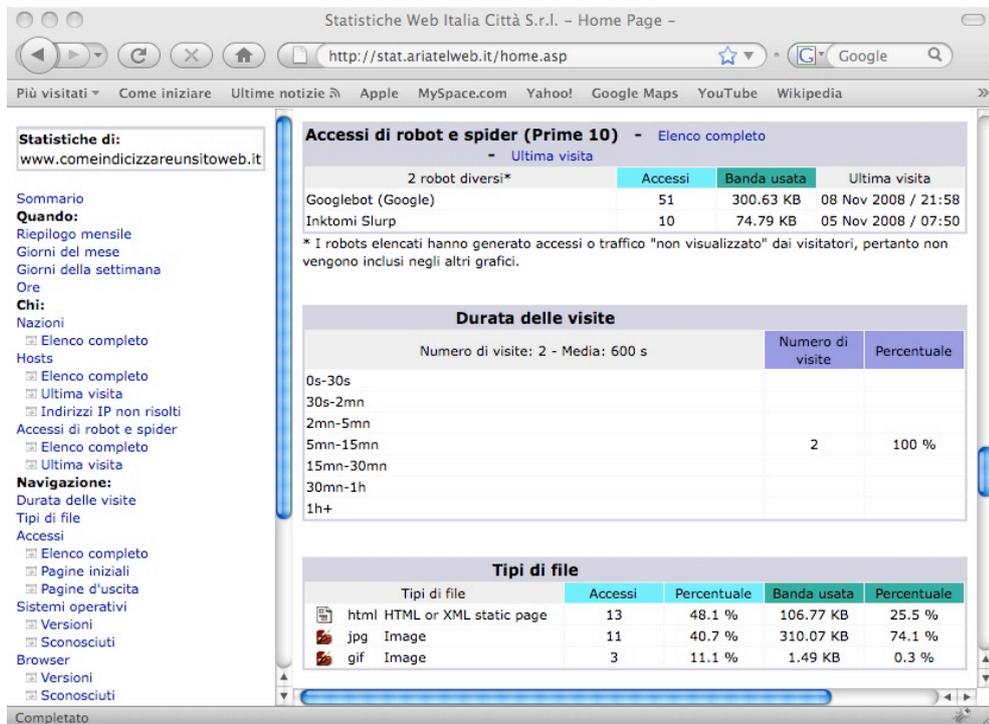
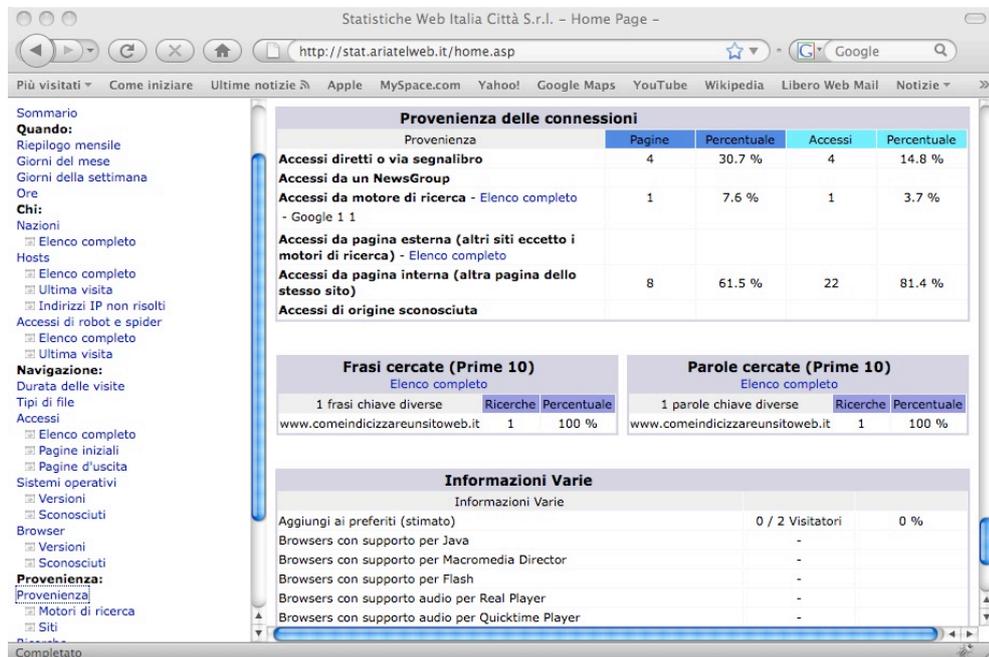


Immagine 4.1: dati statistici riguardanti gli accessi di robot e spider

È inoltre possibile conoscere la provenienza degli accessi, ovvero in che modo un utente è arrivato all'interno del nostro sito web: accessi diretti o via segnalibro, accessi da pagina esterna, accessi da motori di ricerca, etc.

## Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca



Immagini 4.2: dati statistici relativi ai tipi di accessi su stat.ariatelweb.it

In particolare, per quanto riguarda gli accessi tramite ricerche su motori, è fornito anche l'elenco di frasi e parole chiave con cui gli utenti hanno effettuato l'accesso al nostro sito web.

### 4.1.1 DATI STATISTICI SUDDIVISI PER DOMINI WEB

Grazie alla consultazione delle pagine messe a disposizione dell'azienda, è stato possibile ricavare una serie di dati statistici significativi per ogni dominio web.

*Sito web non ottimizzato*

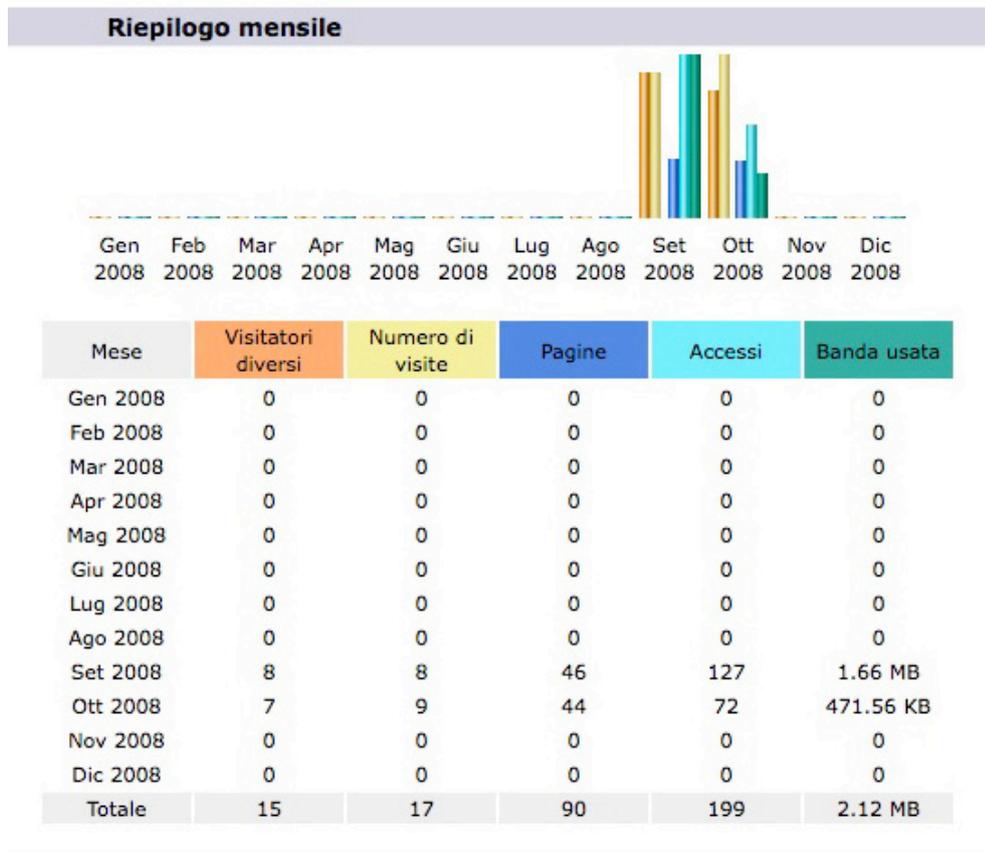


Immagine 4.3: totale visite per lindicizzazione.it

Come si può notare osservando il grafico sopra riportato, il sito [www.lindicizzazione.it](http://www.lindicizzazione.it) è stato visitato da 15 utenti diversi (visite totali 17), per un totale di 90 pagine visitate, 199 accessi e 2.12 MB di banda utilizzata per gli accessi.

Accessi di robot e spider (Prime 10) - <a href="#">Elenco completo</a> -			
Ultima visita			
2 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	108	492.48 KB	31 Ott 2008 / 21:14
MSNBot	1	0	03 Ott 2008 / 11:24

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.4: accessi di robot e spider per lindicizzazione.it

Per quanto riguarda i robot e spider che hanno fatto visita al sito web, sono stati individuati lo spider di Google, GoogleBot, il cui ultimo accesso risale al 31 ottobre, e lo spider di Live Seach, MSNBot, il cui ultimo accesso risale al 3 ottobre. GoogleBot ha eseguito un totale di 108 accessi mentre MSNBot soltanto 1 accesso.

I tipi di accessi riguardanti questo sito web riguardano principalmente gli accessi i diretti o via segnalibro, che interessano 42 pagine, e gli accessi da pagine interne, per quanto riguarda 2 pagine.

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
Accessi diretti o via segnalibro	42	95.4 %	42	58.3 %
Accessi da un NewsGroup				
Accessi da motore di ricerca - <a href="#">Elenco completo</a>				
Accessi da pagina esterna (altri siti eccetto i motori di ricerca) - <a href="#">Elenco completo</a>				
Accessi da pagina interna (altra pagina dello stesso sito)	2	4.5 %	30	41.6 %
Accessi di origine sconosciuta				

Frase cercate (Prime 10)			Parole cercate (Prime 10)		
<a href="#">Elenco completo</a>			<a href="#">Elenco completo</a>		
0 frasi chiave diverse	Ricerche	Percentuale	0 parole chiave diverse	Ricerche	Percentuale

Immagine 4.5: provenienza delle connessioni per lindicizzazione.it

Ovviamente, poiché non ci sono stati accessi tramite motori di ricerca, la lista di frasi e parole chiave risulta vuota.

*Sito web ottimizzato tramite l'utilizzo di meta tag*

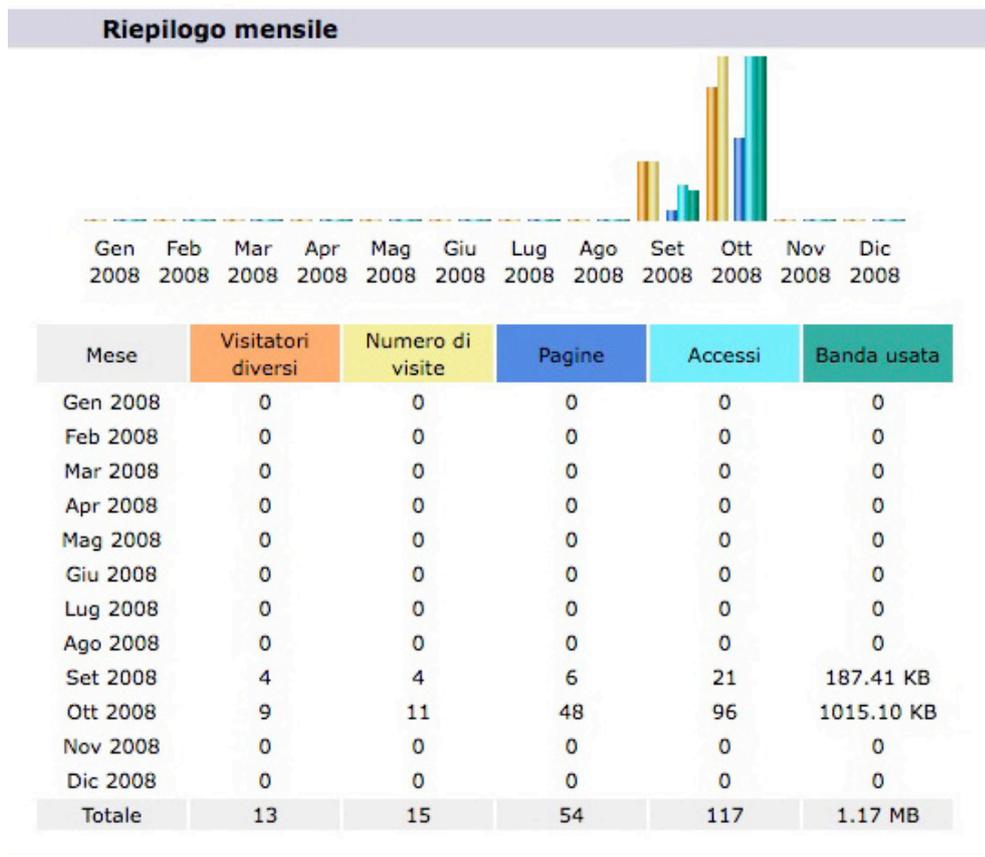


Immagine 4.6: totale visite per [comesiindicizzaunsitoweb.it](http://comesiindicizzaunsitoweb.it)

Osservando il grafico degli accessi, si può notare come le visite ricevute da utenti diversi ammontano a 13 (15 ripetute), per un totale di 54 pagine visitate, 117 accessi e 1.17 MB di banda utilizzata.

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

Accessi di robot e spider (Prime 10) - Elenco completo -			
Ultima visita			
1 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	111	569.86 KB	31 Ott 2008 / 20:39

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.7: accessi di robot e spider per comesiindicizzaunsitoweb.it

Come si può notare, il sito web è stato visitato soltanto da GoogleBot; l'ultimo passaggio dello spider risale al 31 ottobre, per un totale di 111 accessi.

Per quanto riguarda i tipi di accessi, questo dominio è stato raggiunto dagli utenti tramite accessi diretti o via segnalibro per 29 pagine, 7 pagine sono state raggiunte tramite motori di ricerca (in questo caso Google e Ask.com), mentre 12 pagine sono state visitate tramite pagine interne.

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
Accessi diretti o via segnalibro	29	60.4 %	29	30.2 %
Accessi da un NewsGroup				
Accessi da motore di ricerca - Elenco completo	7	14.5 %	7	7.2 %
- Google 6 6				
- Ask Jeeves 1 1				
Accessi da pagina esterna (altri siti eccetto i motori di ricerca) - Elenco completo				
Accessi da pagina interna (altra pagina dello stesso sito)	12	25 %	60	62.5 %
Accessi di origine sconosciuta				

Frase cercate (Prime 10)		
Elenco completo		
6 frasi chiave diverse	Ricerche	Percentuale
motori di ricerca video	1	16.6 %
metadato title	1	16.6 %
metodi di indicizzazione	1	16.6 %
differenza meta description keywords	1	16.6 %
i metadati description keyword	1	16.6 %
metatag pragma	1	16.6 %

Parole cercate (Prime 10)		
Elenco completo		
17 parole chiave diverse	Ricerche	Percentuale
di	2	10.5 %
description	2	10.5 %
keyword	1	5.2 %
metatag	1	5.2 %
ricerca	1	5.2 %
meta	1	5.2 %
keywords	1	5.2 %
indicizzazione	1	5.2 %
pragma	1	5.2 %
differenza	1	5.2 %
Altre parole	7	36.8 %

Immagine 4.8: provenienza delle connessioni per comesiindicizzaunsitoweb.it

Come si può notare dalla lista sopra riportata, per quanto riguarda gli accessi tramite motori di ricerca, la maggior parte degli accessi sono avvenuti tramite

parole chiave inerenti i meta tag e i metadati, ovvero quelle parole che, come vedremo in seguito, sono state maggiormente indicizzate dai motori di ricerca.

*Sito web ottimizzato con l'utilizzo dei BackLink*

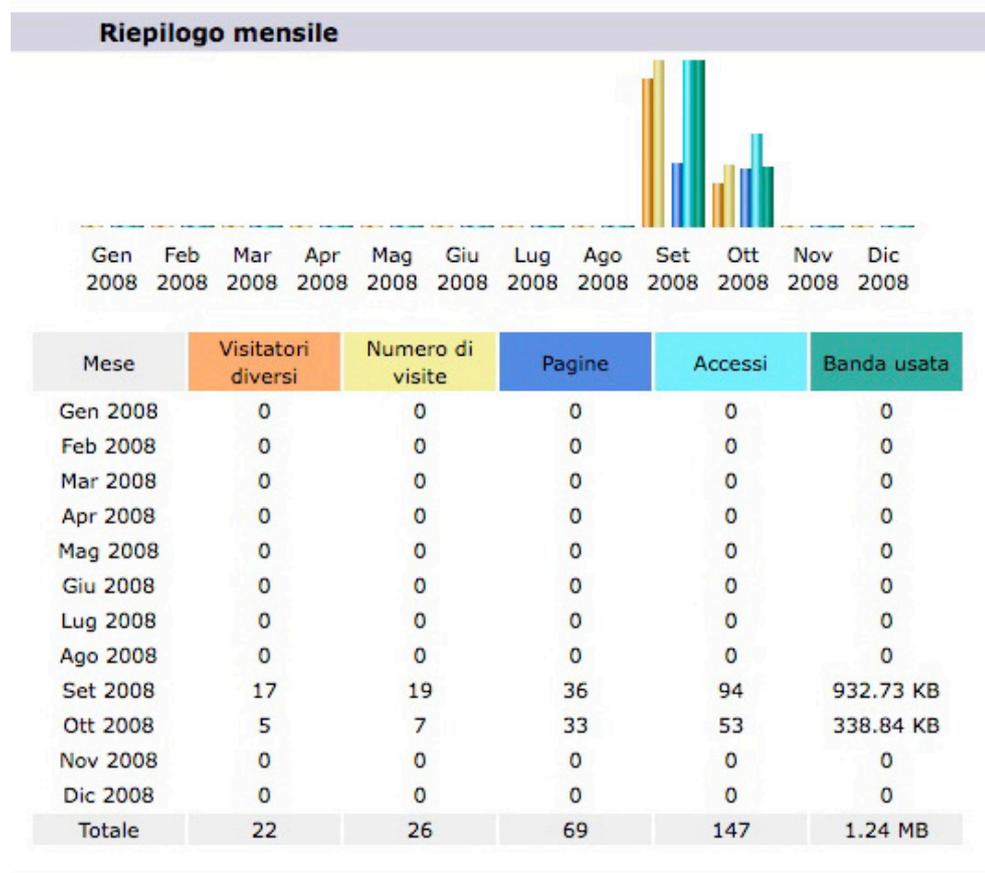


Immagine 4.9: totale visite per ilrankingweb.it

Come si può notare dal grafico sopra riportato, durante i mesi di settembre e ottobre, le visite da parte degli utenti ammontano a 22, per un totale di 69 pagine visitate, 147 accessi e 1.24 MB di banda utilizzata.

Questo dominio web è stato visitato da cinque spider: GoogleBot (Google), Inktomi Slurp e Jeeves (entrambi legati ad Ask.com), MSNBot (Live Search) e GigaBot (spider del motore di ricerca minore Gigablast.com).

Accessi di robot e spider (Prime 10) - <a href="#">Elenco completo</a> -			
Ultima visita			
5 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	133	857.60 KB	31 Ott 2008 / 23:26
Inktomi Slurp	53	358.33 KB	31 Ott 2008 / 03:01
GigaBot	26	197.72 KB	30 Ott 2008 / 21:13
MSNBot	11	70.49 KB	11 Ott 2008 / 01:05
Jeeves	7	57.43 KB	29 Ott 2008 / 03:10

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.10: accessi di robot e spider per ilrankingweb.it

Gli ultimi spider che hanno visitato il nostro sito web sono GoogleBot e Inktomi Slurp, in data 31 ottobre per un totale di 13 accessi per il primo spider e 53 per il secondo. GigaBot ha effettuato il suo ultimo accesso in data 30 ottobre, raggiungendo la quota di 26 accessi, MSNBot in data 11 ottobre con 11 accessi e Jeeves in data 29 ottobre con 7 accessi.

Le visite al nostro sito web sono state eseguite tramite accesso diretto o via segnalibro per quanto riguarda 31 pagine; due pagine invece sono state visitate tramite ricerca sul motore Google.

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
<b>Accessi diretti o via segnalibro</b>	31	93.9 %	31	58.4 %
<b>Accessi da un NewsGroup</b>				
<b>Accessi da motore di ricerca</b> - <a href="#">Elenco completo</a>	2	6 %	2	3.7 %
- Google 2 2				
<b>Accessi da pagina esterna (altri siti eccetto i motori di ricerca)</b> - <a href="#">Elenco completo</a>			5	9.4 %
- http://74.125.39.104/search 0 5				
<b>Accessi da pagina interna (altra pagina dello stesso sito)</b>			15	28.3 %
<b>Accessi di origine sconosciuta</b>				

Frase cercate (Prime 10)		
<a href="#">Elenco completo</a>		
2 frasi chiave diverse	Ricerche	Percentuale
portale virgilio storia	1	50 %
meta pragma	1	50 %

Parole cercate (Prime 10)		
<a href="#">Elenco completo</a>		
5 parole chiave diverse	Ricerche	Percentuale
pragma	1	20 %
virgilio	1	20 %
portale	1	20 %
storia	1	20 %
meta	1	20 %

Immagine 4.11: provenienza delle connessioni per ilrankingweb.it

Le frasi attraverso le quali gli utenti hanno effettuato l'accesso al dominio web, come si può notare, sono due: "portale Virgilio storia" e "meta pragma" (in effetti, per quest'ultima parola-chiave, il sito risulta indicizzato al secondo posto per Google).

*Sito web ottimizzato con l'utilizzo di Sitemap*

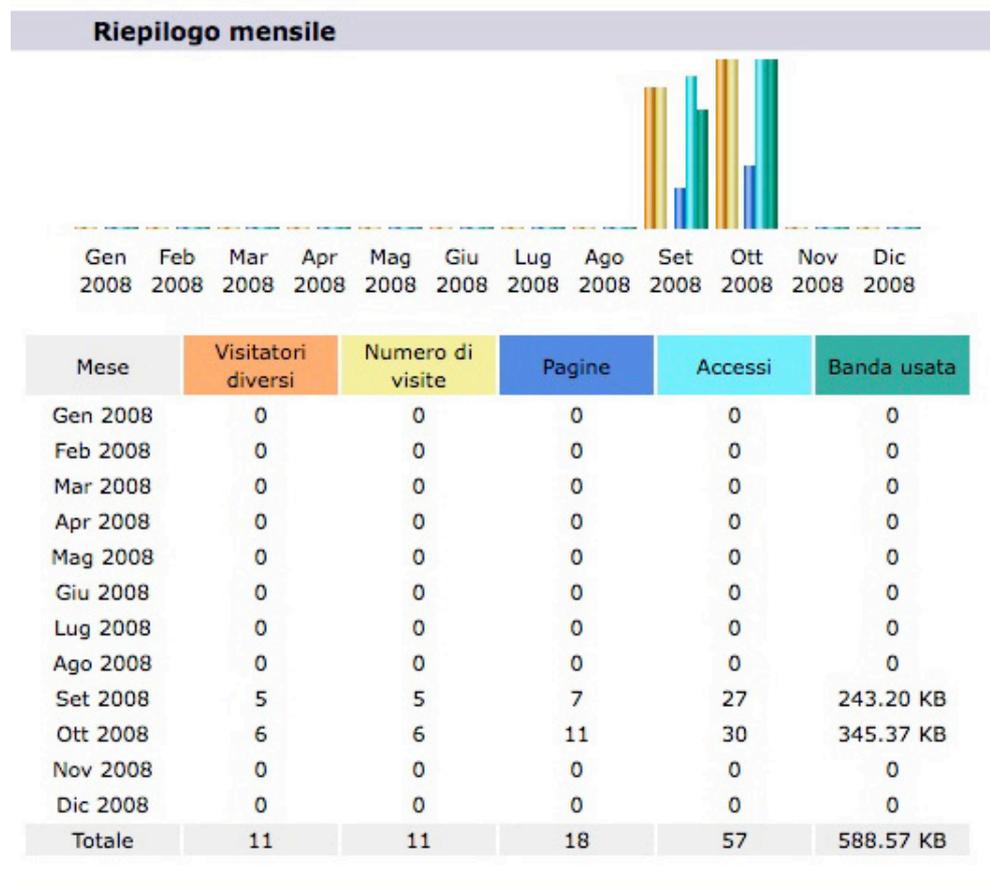


Immagine 4.12: totale visite per indicizzazioneweb.it

Il sito web preso in esame ha ricevuto, nelle cinque settimane di messa online, 11 visite, per un totale di 18 pagine visitate, 57 accessi e un 58.57 KB di banda utilizzata per la connessione.

www.indicizzazioneweb.it è stato visitato, in questo intervallo di tempo, da due spider: GoogleBot, che ha effettuato il suo ultimo passaggio in data 31 ottobre, raggiungendo i 258 accessi, e uno spider sconosciuto, indicato con "Unknown robot (identified by 'crawl')", che ha effettuato la sua ultima visita in data 2 ottobre raggiungendo un numero di accessi pari a 24.

Accessi di robot e spider (Prime 10) - Elenco completo -			
Ultima visita			
2 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	258	1.22 MB	31 Ott 2008 / 22:51
Unknown robot (identified by 'crawl')	24	170.29 KB	02 Ott 2008 / 20:51

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.13: accessi di robot e spider per indicizzazioneweb.it

Come si può notare dalla tabella seguente, gli accessi effettuati a questo sito web sono avvenuti tramite accesso diretto o via segnalibro (6 accessi) e tramite accessi da pagina interna (5 accessi). Poiché non sono avvenuti accessi tramite ricerche di parole chiave, la lista delle frasi e delle parole legate a questo sito web, risulta vuota.

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
Accessi diretti o via segnalibro	6	54.5 %	6	20 %
Accessi da un NewsGroup				
Accessi da motore di ricerca - Elenco completo				
Accessi da pagina esterna (altri siti eccetto i motori di ricerca) - Elenco completo				
Accessi da pagina interna (altra pagina dello stesso sito)	5	45.4 %	24	80 %
Accessi di origine sconosciuta				

Frase cercate (Prime 10)			Parole cercate (Prime 10)		
Elenco completo			Elenco completo		
0 frasi chiave diverse	Ricerche	Percentuale	0 parole chiave diverse	Ricerche	Percentuale

Immagine 4.14: provenienza delle connessioni per indicizzazioneweb.it

*Sito web ottimizzato con l'utilizzo dei tag HTML*

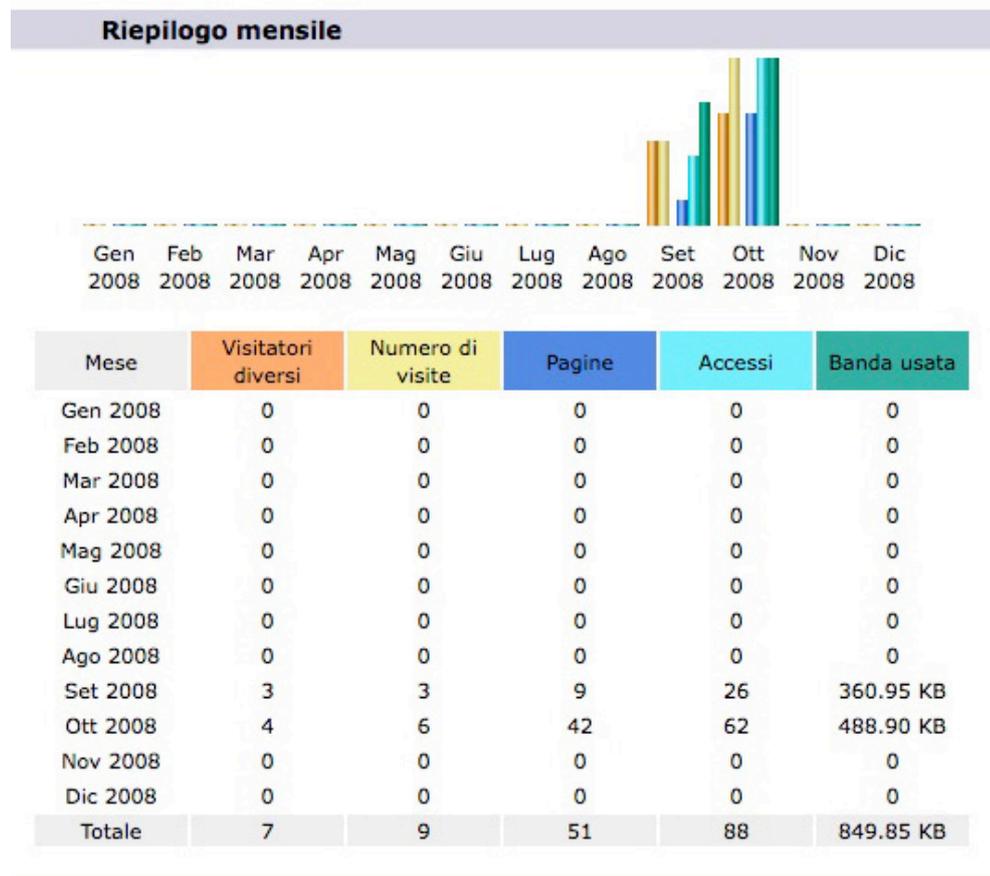


Immagine 4.15: totale accessi per motoriealgoritmidiricerca.it

Il sito web [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it) ha ricevuto 7 visite durante l'intervallo di tempo preso in esame, per un totale di 51 pagine visitate, 88 accessi e 849.85 KB di banda utilizzata.

Questo dominio web è stato visitato nel corso della sperimentazione da un solo spider, GoogleBot di Google, che ha effettuato il suo ultimo accesso in data 31 ottobre, raggiungendo la quota di 111 accessi totali.

Accessi di robot e spider (Prime 10) - Elenco completo -			
Ultima visita			
1 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	111	497.01 KB	31 Ott 2008 / 19:05

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.16: accessi di robot e spider per motoriealgorithmidricerca.it

Per quanto riguarda la provenienza delle connessioni, 33 pagine sono state visualizzate tramite accessi diretti o via segnalibro; 9 pagine invece hanno ricevuto visite tramite accessi da pagina interna.

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
Accessi diretti o via segnalibro	33	78.5 %	33	53.2 %
Accessi da un NewsGroup				
Accessi da motore di ricerca - Elenco completo				
Accessi da pagina esterna (altri siti eccetto i motori di ricerca) - Elenco completo				
Accessi da pagina interna (altra pagina dello stesso sito)	9	21.4 %	29	46.7 %
Accessi di origine sconosciuta				

Frase cercate (Prime 10)			Parole cercate (Prime 10)		
Elenco completo			Elenco completo		
0 frasi chiave diverse	Ricerche	Percentuale	0 parole chiave diverse	Ricerche	Percentuale

Immagine 4.17: provenienza delle connessioni per motoriealgorithmidricerca.it

Come si può ben notare osservando la tabella sopra riportata, poiché non vi sono stati accessi tramite motori di ricerca, le tabelle riguardanti le frasi e le parole chiave, con cui uno o più utenti possono aver avuto accesso alla pagina web, risultano vuote.

*Sito web ottimizzato con tutte le tecniche applicate singolarmente negli altri domini*

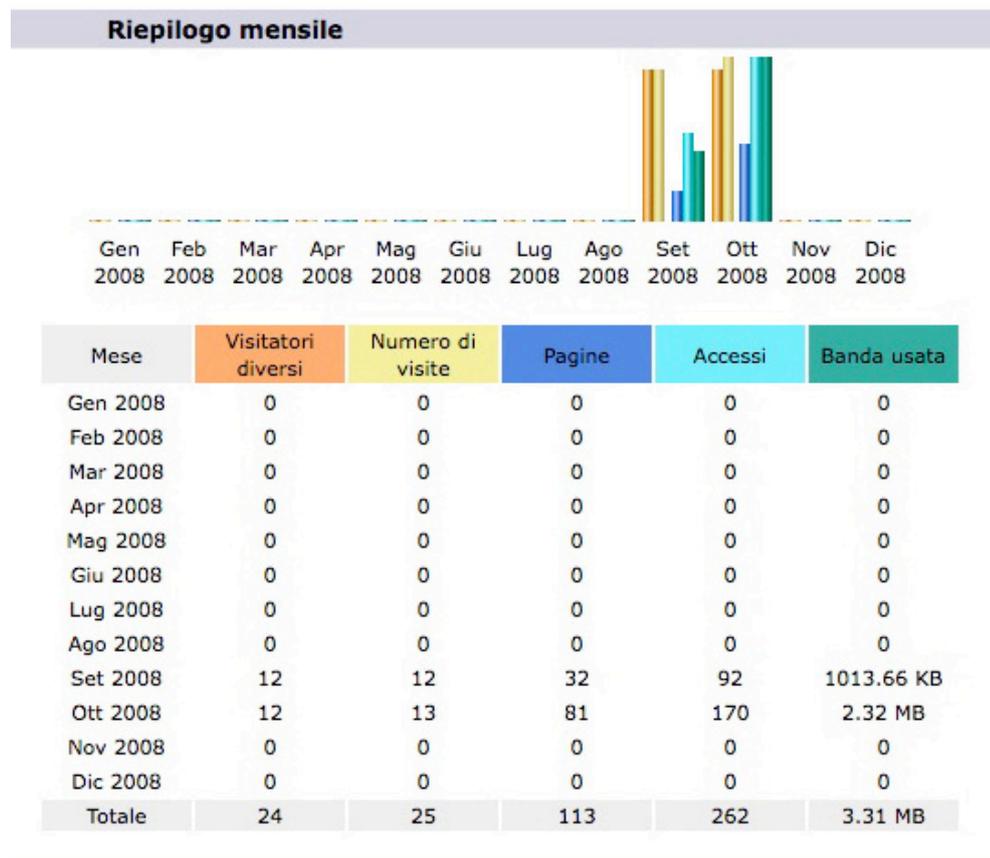


Immagine 4.18: totale visite per comeindicizzareunsitoweb.it

Il sito web [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it), il dominio web sul quale sono state applicate tutte le singole tecniche di ottimizzazione viste singolarmente nei siti web analizzati in precedenza, ha ricevuto 24 visite, per un totale di 113 pagine visitate, 262 accessi e 3.31 MB di banda utilizzata.

Il seguente sito web è stato visitato da tre spider: GoogleBot, che ha eseguito l'ultima visita in data 31 ottobre, per un totale di 240 accessi; uno spider non identificato, che ha effettuato l'ultima visita in data 2 ottobre, per un totale di 24 accessi e Inktomi Slurp, che ha effettuato l'ultimo accesso in data 31 ottobre per un totale di 29 accessi.

Accessi di robot e spider (Prime 10) - <a href="#">Elenco completo</a> -			
Ultima visita			
3 robot diversi*	Accessi	Banda usata	Ultima visita
Googlebot (Google)	240	1.31 MB	31 Ott 2008 / 20:07
Inktomi Slurp	29	217.37 KB	31 Ott 2008 / 21:59
Unknown robot (identified by 'crawl')	24	200.35 KB	02 Ott 2008 / 20:38

\* I robots elencati hanno generato accessi o traffico "non visualizzato" dai visitatori, pertanto non vengono inclusi negli altri grafici.

Immagine 4.19: provenienza delle connessioni per [comeindicizzareunsitoweb.it](http://comeindicizzareunsitoweb.it)

Le visite al sito web sono state eseguite tramite accessi diretti o via segnalibro per quanto riguarda 15 pagine, attraverso accessi da pagina interna per 65 pagine e tramite ricerca su motore per una sola pagina.

Provenienza delle connessioni				
Provenienza	Pagine	Percentuale	Accessi	Percentuale
Accessi diretti o via segnalibro	15	18.5 %	15	8.8 %
Accessi da un NewsGroup				
Accessi da motore di ricerca - <a href="#">Elenco completo</a>	1	1.2 %	1	0.5 %
- Google 1 1				
Accessi da pagina esterna (altri siti eccetto i motori di ricerca) - <a href="#">Elenco completo</a>				
Accessi da pagina interna (altra pagina dello stesso sito)	65	80.2 %	154	90.5 %
Accessi di origine sconosciuta				

Frase cercate (Prime 10)			
<a href="#">Elenco completo</a>			
1 frasi chiave diverse	Ricerche	Percentuale	
alt linguaggio w3c	1	100 %	

Parole cercate (Prime 10)			
<a href="#">Elenco completo</a>			
3 parole chiave diverse	Ricerche	Percentuale	
alt	1	33.3 %	
linguaggio	1	33.3 %	
w3c	1	33.3 %	

Immagine 4.20: provenienza connessioni per comeindicizzareunsitoweb.it

Nel caso di accesso tramite ricerca su motore, la frase ricercata è stata “alt linguaggio w3c”.

#### 4.1.2 SINTESI E COMPARAZIONE DEI DATI STATISTICI

Procedendo con un'analisi comparata dei dati statistici per i singoli siti web, si può notare come il dominio che ha ricevuto il minor numero di visite (dove con “visite” s'intende il numero di utenti diversi che hanno eseguito l'accesso ad almeno una pagina web del nostro sito) è stato [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it), ovvero il sito web indicizzato tramite l'utilizzo dei tag HTML. Come vedremo in seguito, in effetti, questa tecnica di ottimizzazione ha portato solo un debole miglioramento della posizione di alcune parole chiave.

Viceversa, il dominio web che ha ricevuto il maggior numero di visite risulta [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it). Si può dire che questo dato fosse prevedibile e coerente con l'analisi in quanto, il sito che ha ricevuto il maggior numero di visite, è proprio quello a cui sono state applicate tutte le tecniche di ottimizzazione viste singolarmente nei restanti domini, e quindi quello più esposto all'attenzione di motori di ricerca e utenti.

Analizzando i dati riguardanti le visite di spider e robot dei motori di ricerca, emerge che l'unico spider che ha visitato tutti e sei i domini creati è GoogleBot, il robot del motore di ricerca Google. Inoltre, per ogni dominio web, lo spider di Google è quello che ha compiuto il maggior numero di visite ai siti. Nel caso di domini web come [www.comesiindicizzaunsitoweb.it](http://www.comesiindicizzaunsitoweb.it) e [www.motoriealgoritmidiricerca.it](http://www.motoriealgoritmidiricerca.it), GoogleBot si è inoltre rivelato l'unico spider che ha effettuato visite.

MSNBot invece, lo spider del motore di ricerca Live Search, ha visitato soltanto due siti su sei: [www.lindicizzazione.it](http://www.lindicizzazione.it) e [www.ilrankingweb.it](http://www.ilrankingweb.it)

Anche Inktomi slurp, spider del motore di ricerca Ask.com, ha eseguito visite in due siti web: [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it) e [www.ilrankingweb.it](http://www.ilrankingweb.it).

Questo ultimo sito web è quello che ha ricevuto visite da un maggior numero di spider; oltre a GoogleBot, MSNBot e Inktomi Slurp, è stato infatti visitato dallo spider Jeeves, anch'esso di Ask.com, e da GigaBot, robot del motore di ricerca secondario GigaBlast.com

Inoltre alcuni siti web, in particolare [www.indicizzazioneweb.it](http://www.indicizzazioneweb.it) e [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it), sono stati visitati da uno spider non identificato, indicato nelle statistiche con "Unknow robot (identified by 'crawl')".

In conclusione, le visite ricevute dai siti web raccolgono gli spider di quasi tutti i motori di ricerca presi in considerazione durante l'esperimento. Anche se il solo che ha effettuato almeno un accesso per ogni sito web è stato lo spider di Google, è possibile conteggiare almeno due visite da parte di ogni motore di ricerca. All'interno delle statistiche non si trova indicazione del passaggio degli spider di Arianna o Virgilio poiché, come già spiegato in precedenza, questi due motori di ricerca italiano si appoggiano ai risultati forniti da Google.

Analizzando i dati, emerge in ogni caso la mancanza di visite da parte del robot del sesto motore di ricerca da noi analizzato: Yahoo!.

Se si osservano i dati relativi ai tipi di accessi che hanno portato gli utenti a visitare i siti web pubblicati, si può notare come in tutti i casi la maggior parte

degli accessi sono avvenuti tramite accesso diretto o via segnalibro, ovvero tramite la digitazione diretta dell'indirizzo all'interno della barra del browser. Questo dato è spiegabile in quanto, all'interno del numero fornito dalle statistiche, sono inclusi anche gli accessi che io stessa ho eseguito per controllare la corretta visualizzazione delle pagine una volta che queste ultime sono state rese visibili online. Successivo a questo dato, abbastanza alte sono le visite avvenute attraverso pagina interna, ovvero tramite link interni al sito stesso.

L'accesso ai siti web tramite sottomissione di parole chiave ai motori di ricerca è avvenuto nel 50% dei casi, ovvero per tre siti su sei: [www.ilrankingweb.it](http://www.ilrankingweb.it), [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it), [www.comeiindicizzaunsitoweb.it](http://www.comeiindicizzaunsitoweb.it). Quest'ultimo risulta essere il sito web visitato tramite il maggior numero di frasi digitate all'interno di un motore di ricerca; è inoltre l'unico che ha ricevuto un accesso tramite il motore di ricerca Ask.com.

Analizzando le frasi e le parole chiave che si sono rivelate vincenti per l'accesso ai nostri domini, si può notare come la maggior parte riguardano i meta tag e i metadati, oltre ad una serie di ricerche tramite le frasi "metodi di indicizzazione", "alt linguaggio w3c" o "motori di ricerca video".

La significativa presenza di parole come "metadati", "meta title", "meta description", può essere spiegata col fatto che queste sono parole tecniche, legate comunque al mondo del web marketing. Altre parole chiave utilizzate per le analisi, come per esempio "parole chiave", "motori di ricerca", "ottimizzazione siti web", possono invece essere impiegate più frequentemente anche da coloro che non si occupano specificatamente di indicizzazione.

Come sarà visto nel dettaglio in seguito, i siti web che sono stati effettivamente indicizzati, hanno avuto maggiore ottimizzazione proprio per quel gruppo di parole chiave legate ai meta tag.

## **4.2 RACCOLTA E ANALISI DEI DATI RELATIVI AL POSIZIONAMENTO**

Trascorse le tre settimane stabilite, si è tornati ad utilizzare il programma Accurate Monitor for Search Engine. Attraverso l'utilizzo di quest'ultimo, sono stati ottenuti i dati relativi al posizionamento delle parole chiave stabilite in precedenza per i diversi siti web, sui sei motori di ricerca presi in considerazione per l'analisi.

Attraverso l'utilizzo di questo programma sono state quindi effettuate, per ogni dominio web, quattro scansioni: per Google, per Yahoo!, per Live Search e per Ask.com. Per i due restanti motori di ricerca, ovvero i portali italiani Arianna e Virgilio, non è stato necessario eseguire alcuna scansione poiché i risultati presentati sembrano essere i medesimi del motore di ricerca Google.

Effettuate le singole scansioni, sono stati costruiti dei fogli di lavoro in Excel per avere una visione d'insieme sui dati ottenuti.

Dai risultati ottenuti si è potuto notare come i domini web siano stati indicizzati soltanto da Google, e di conseguenza da Virgilio e Arianna, e, in misura nettamente minore, da Ask.com.

### **4.2.1 GOOGLE**

#### *Sito web non ottimizzato*

Per quanto riguarda il sito non modificato da nessuna tecnica SEO, è stata effettuata la scansione su 775 posizioni (il massimo consentito dal programma utilizzato):

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

	<b>Sito non ottimizzato</b>
<b>Parole chiave</b>	
ask.com	>775
motori di ricerca	>775
algoritmi di ricerca	>775
motori e algoritmi di ricerca	>775
alt sulle immagini	>775
metadati	>775
linguaggio HTML	>775
arianna	>775
backlink	>775
posizionamento siti web	>775
metodi di indicizzazione	>775
link popularity	>775
google	>775
pagerank	>775
tecniche seo	>775
ottimizzazione siti web	>775
tag HTML	>775
keywords density	>775
parole-chiave	>775
meta tag	>775
meta title	251
meta description	394
meta keywords	123
meta robots	262
meta pragma	54
meta revisit-after	124
directory	>775
virgilio	>775
yahoo	>775

algoritmo edison	>775
algoritmo teoma	>775
expertrank	>775

Come si può notare dalla tabella sopra riportata, le uniche parole che hanno subito un miglioramento di posizione sono quelle che riguardano i meta tag, ovvero: meta title, meta description, meta keywords, meta robots, meta pragma, meta revisit-after. Come spiegato in precedenza, probabilmente il miglioramento di posizioni per queste parole chiave è da spiegarsi in relazione al loro grado di tecnicismo, che le rende meno utilizzate rispetto ad altre d'uso comune.

La necessità di compiere una scansione su un così alto numero di posizioni è stata dettata dal fatto che occorre numeri significativi sui quali fare un paragone rispetto alle scansioni effettuate per i restanti domini web.

I successivi domini web sono stati invece scansionati su 300 posizioni.

*Sito web ottimizzato tramite l'utilizzo dei meta tag*

	<b>Sito non ottimizzato</b>	<b>Meta tag</b>
<b>Parole chiave</b>		
ask.com	>775	>300
motori di ricerca	>775	>300
algoritmi di ricerca	>775	>300
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	>300
metadati	>775	>300
linguaggio HTML	>775	>300
arianna	>775	>300
backlink	>775	>300
posizionamento siti web	>775	>300

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	>300
parole-chiave	>775	>300
meta tag	>775	>300
meta title		251 29 <b>(+222)</b>
meta description		394 96 <b>(+298)</b>
meta keywords		123 51 <b>(+72)</b>
meta robots		262 210 <b>(+52)</b>
meta pragma		54 37 <b>(+17)</b>
meta revisit-after		124 117 <b>(+7)</b>
directory	>775	>300
virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	22 <b>(+&gt;763)</b>
algoritmo teoma	>775	1 <b>(+&gt;774)</b>
expertrank	>775	34 <b>(+&gt;741)</b>

La seguente tabella riporta, accanto alla lista delle parole-chiave, il confronto di posizionamento tra il sito web [www.lindicizzazione.it](http://www.lindicizzazione.it), di cui si è precedentemente parlato, e il sito [www.comesiindicizzaunsitoweb.it](http://www.comesiindicizzaunsitoweb.it), scansionato appunto su 300 posizioni. Nella tabella i numeri rappresentati in rosso indicano, per quelle parole-chiave che hanno subito un incremento di posizioni, il numero di posizioni “guadagnate” rispetto al sito che non ha subito alcun tipo di modifica. Come si può notare, le parole che sono state ottimizzate riguardano il gruppo dei meta tag, così com’era avvenuto anche per [www.lindicizzazione.it](http://www.lindicizzazione.it) e le parole-chiavi algoritmo Edison, algoritmo teoma ed expertrank.

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

*Sito web ottimizzato tramite i BackLink*

	<b>Sito non ottimizzato</b>	<b>BackLink</b>
<b>Parole chiave</b>		
ask.com	>775	>300
motori di ricerca	>775	>300
algoritmi di ricerca	>775	>300
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	>300
metadati	>775	>300
linguaggio HTML	>775	>300
arianna	>775	>300
backlink	>775	>300
posizionamento siti web	>775	>300
metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	44 ( <b>+&gt;731</b> )
parole-chiave	>775	>300
meta tag	>775	>300
meta title	251	9 ( <b>+&gt;766</b> )
meta description	394	>300
meta keywords	123	65 ( <b>+&gt;710</b> )
meta robots	262	49 ( <b>+&gt;726</b> )
meta pragma	54	2 ( <b>+&gt;773</b> )
meta revisit-after	124	5 ( <b>+&gt;770</b> )

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

directory	>775	>300
virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	44 ( <b>+&gt;731</b> )
algoritmo teoma	>775	>300
expertrank	>775	>300

Questi sono i risultati ottenuti per il dominio web ottimizzato con l'uso dei cosiddetti BackLink. Come si può notare, anche in questo caso le parole-chiave che hanno subito un incremento di posizione sono quelle del gruppo dei meta tag, ad esclusione della parola-chiave "meta description" che in questo caso non è stata indicizzata, almeno non al di sopra della trecentesima posizione. In questo dominio sono inoltre state indicizzate le parole algoritmo Edison e keywords density, che hanno entrambe guadagnato più di 731 posizioni. Da notare in particolare le parole-chiave meta title, meta pragma e meta revisit-after che hanno ottenuto una posizione tra i primi dieci risultati (meta pragma addirittura la seconda posizione), ovvero un posizionamento nella prima pagina di risultati del motore di ricerca Google.

*Sito web ottimizzato tramite l'aggiunta di sitemap*

	<b>Sito non ottimizzato</b>	<b>Sitemap</b>
<b>Parole chiave</b>		
ask.com	>775	>300
motori di ricerca	>775	>300
algoritmi di ricerca	>775	>300
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	>300
metadati	>775	>300
linguaggio HTML	>775	>300
arianna	>775	>300

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

backlink	>775	>300
posizionamento siti web	>775	>300
metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	>300
parole-chiave	>775	>300
meta tag	>775	>300
meta title	251	>300
meta description	394	>300
meta keywords	123	>300
meta robots	262	>300
meta pragma	54	>300
meta revisit-after	124	>300
directory	>775	>300
virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	>300
algoritmo teoma	>775	>300
expertrank	>775	>300

Il dominio web [www.indicizzazioneweb.it](http://www.indicizzazioneweb.it), ottimizzato tramite l'uso di una sitemap, mostra come questa tecnica nel nostro caso non abbia portato ai risultati sperati, ovvero un incremento di posizioni per le parole-chiave elencate. Si può addirittura notare un peggioramento per la posizione di alcune di queste: per esempio meta title ha perso più di 49 posizioni, meta keyword più di 177, meta robots più di 38, meta pragma più di 256 e infine meta revisit-after più di 176.

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

*Sito web ottimizzato tramite l'utilizzo di tag HTML*

	<b>Sito non ottimizzato</b>	<b>Linguaggio HTML</b>
<b>Parole chiave</b>		
ask.com	>775	>300
motori di ricerca	>775	>300
algoritmi di ricerca	>775	141 ( <b>+&gt;634</b> )
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	>300
metadati	>775	>300
linguaggio HTML	>775	>300
arianna	>775	>300
backlink	>775	>300
posizionamento siti web	>775	>300
metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	>300
parole-chiave	>775	>300
meta tag	>775	>300
meta title	251	>300
meta description	394	>300
meta keywords	123	>300
meta robots	262	>300
meta pragma	54	2 ( <b>+&gt;52</b> )
meta revisit-after	124	135 ( <b>-9</b> )
directory	>775	>300

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	>300
algoritmo teoma	>775	>300
expertrank	>775	>300

Questo dominio web è stato ottimizzato tramite l'utilizzo dei tag HTML; esaminando la tabella sopra riportata, si può notare come le parole che hanno subito un incremento nella lista di risultati dei motori di ricerca siano poche rispetto a quelle indicizzate tramite i meta tag e i BackLink. Esaminando i dati ricavati, le frasi indicizzate risultano essere soltanto due: "algoritmi di ricerca" (probabilmente per il fatto che la parola-chiave è presente nel nome del dominio stesso) e "meta pragma" (salito addirittura in seconda posizione). La frase "meta revisit-after" invece ha peggiorato la sua posizione di 9 posti rispetto al sito web [www.lindicizzazione.it](http://www.lindicizzazione.it).

*Sito web ottimizzato tramite l'utilizzo di tutte le tecniche SEO viste in precedenza*

	<b>Sito non ottimizzato</b>	<b>Sito ottimizzato con tutte le tecniche</b>
<b>Parole chiave</b>		
ask.com	>775	179 <b>(+&gt;596)</b>
motori di ricerca	>775	>300
algoritmi di ricerca	>775	>300
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	76 <b>(+&gt;699)</b>
metadati	>775	>300

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

linguaggio HTML	>775	>300
arianna	>775	>300
backlink	>775	>300
posizionamento siti web	>775	>300
metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	>300
parole-chiave	>775	>300
meta tag	>775	>300
meta title	251	28 <b>(+223)</b>
meta description	394	82 <b>(+312)</b>
meta keywords	123	50 <b>(+123)</b>
meta robots	262	24 <b>(+238)</b>
meta pragma	54	39 <b>(+15)</b>
meta revisit-after	124	10 <b>(+114)</b>
directory	>775	>300
virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	9 <b>(+&gt;766)</b>

Valutazione dell'efficacia di tecniche di web authoring per il miglioramento del ranking nei motori di ricerca

algoritmo teoma	>775	9 <b>(+&gt;766)</b>
expertrank	>775	33 <b>(+&gt;742)</b>

Il sesto e ultimo dominio web, [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it), è stato ottimizzato grazie all'utilizzo di tutte le tecniche viste singolarmente nei precedenti siti internet. Come si può notare, questo sito internet è stato ottimizzato per le parole chiave del gruppo dei meta tag e le seguenti parole chiave: ask.com, alt sulle immagini, algoritmo Edison, algoritmo teoma ed expertrank.

#### 4.2.2 COMPARAZIONE DEI DATI OTTENUTI PER IL MOTORE DI RICERCA GOOGLE

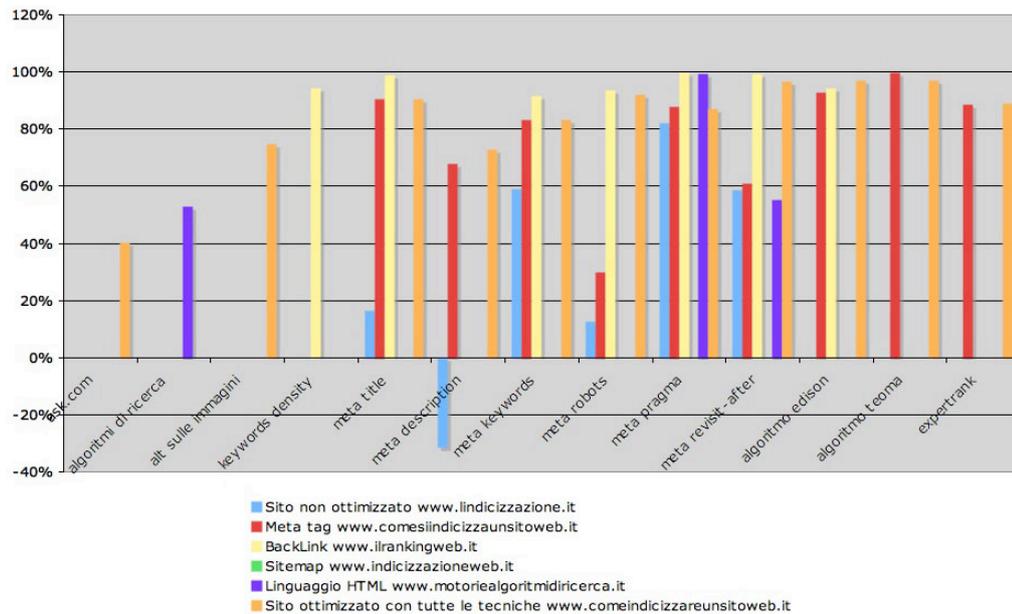


Immagine 4.21: grafico comparativo dell'indicizzazione delle parole chiave nei diversi domini

Il grafico sopra riportato illustra l'incremento di indicizzazione delle parole chiave per i diversi siti web, e quindi per le diverse tecniche SEO.

Naturalmente, nel grafico sono stati inseriti i dati riguardanti quelle parole-chiave che per almeno un dominio hanno subito un incremento nel posizionamento all'interno della SERP di Google.

I dati sono riportati in percentuale perché è importante capire non tanto il valore assoluto di posizioni guadagnate dalla parola-chiave, ma l'influenza che

quest'incremento ha sulla posizione che le stesse parole hanno nel sito non modificato da tecniche di indicizzazione.

Per avere una più chiara visione d'insieme, è stato creato un ulteriore grafico, che illustra la media, ancora una volta in percentuale, degli incrementi ottenuti dai siti web per tutte le parole chiave che sono state indicizzate:

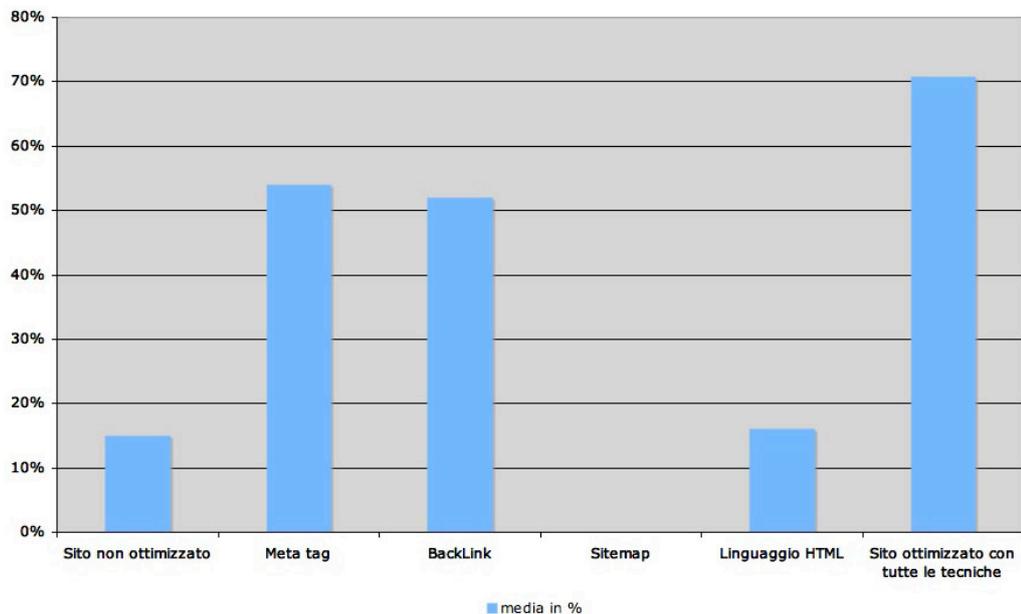


Immagine 4.22: grafico comparativo della media di incrementi ottenuti dai diversi domini

Come si può osservare dal grafico sopra riportato, il sito web che ha registrato un incremento medio maggiore, pari al 70,79%, è quello modificato tramite tutte le tecniche SEO analizzate. Successivamente, con una media vicina al 50%, si trova il sito ottimizzato tramite l'utilizzo di meta tag e quello in cui sono stati utilizzati i BackLink. Questo dato è quindi coerente con l'ipotesi fatta all'inizio della sperimentazione, ovvero il fatto che ottimizzare un dominio

tramite l'uso combinato di diverse tecniche permette di guadagnare maggiori posizioni rispetto ad un sito ottimizzato con una solo metodo.

#### 4.2.3 ASK.COM

	<b>Sito non ottimizzato</b>	<b>BackLink</b>
<b>Parole chiave</b>		
ask.com	>775	>300
motori di ricerca	>775	>300
algoritmi di ricerca	>775	>300
motori e algoritmi di ricerca	>775	>300
alt sulle immagini	>775	>300
metadati	>775	>300
linguaggio HTML	>775	>300
arianna	>775	>300
backlink	>775	>300
posizionamento siti web	>775	>300
metodi di indicizzazione	>775	>300
link popularity	>775	>300
google	>775	>300
pagerank	>775	>300
tecniche seo	>775	>300
ottimizzazione siti web	>775	>300
tag HTML	>775	>300
keywords density	>775	>300
parole-chiave	>775	>300
meta tag	>775	>300
meta title	>775	6 ( <b>+&gt;769</b> )
meta description	>775	>300
meta keywords	>775	>300

meta robots	>775	>300
meta pragma	>775	3 ( <b>+&gt;772</b> )
meta revisit-after	>775	7 ( <b>+&gt;768</b> )
directory	>775	>300
virgilio	>775	>300
yahoo	>775	>300
algoritmo edison	>775	7 ( <b>+&gt;768</b> )
algoritmo teoma	>775	>300
expertrank	>775	14 ( <b>+&gt;761</b> )

Il motore di ricerca Ask.com è l'unico, insieme a Google, su cui è stato possibile ottimizzare i contenuti dei nostri siti web per determinate parole chiave. In ogni caso, è necessario sottolineare il fatto che il numero delle parole indicizzate da questo motore di ricerca è nettamente inferiore rispetto a quelle indicizzate da Google; inoltre l'ottimizzazione è avvenuta per un solo sito web, [www.ilrankingweb.it](http://www.ilrankingweb.it).

Riportiamo quindi i dati per questo solo sito web, comparandoli con la scansione effettuata su [www.lindicizzazione.it](http://www.lindicizzazione.it), costante punto di riferimento per la ricerca.

#### 4.2.4 LIVE SEARCH E YAHOO!

I motori di ricerca Live Search e Yahoo! sono gli unici per i quali, effettuando una scansione su 775 posizioni per [www.lindicizzazione.it](http://www.lindicizzazione.it) e su 300 posizioni per i restanti siti web, non sono stati registrati dati riguardanti l'indicizzazione di parole chiave.

#### 4.2.5 ARIANNA E VIRGILIO

Come spiegato ampiamente in precedenza, questi due motori di ricerca italiani si appoggiano, per le ricerche effettuate al loro interno, ai risultati forniti da Google.

Non sono stati quindi creati grafici che mostrassero i dati ottenuti per l'ottimizzazione delle parole chiave in questi motori di ricerca semplicemente perché i risultati sarebbero del tutto simili a quelli raccolti per il motore di ricerca statunitense.

L'assoluta coerenza tra i due motori di ricerca italiani e Google, per quanto riguarda i dati sul posizionamento, è stata verificata tramite una serie di ricerche sui crawler Virgilio e Arianna. Questi test hanno infatti evidenziato una sostanziale similarità di posizionamento nei tre motori di ricerca per i domini web creati.

## **CONCLUSIONI**

### **COMPORTAMENTO DEI MOTORI DI RICERCA**

Dall'analisi condotta, è possibile affermare che Google si è dimostrato il motore di ricerca che meglio risponde alle esigenze di un sito web dinamico, i cui contenuti variano velocemente nel corso del tempo. Tale affermazione si può dedurre dal fatto che, prima di tutto, questo motore di ricerca è l'unico, tra i sei considerati per l'analisi, che ha inviato il suo spider in tutti i siti web pubblicati. Soprattutto, in tutti i casi esaminati, lo spider GoogleBot è quello che ha eseguito il maggior numero di accessi; questi inoltre sono nettamente maggiori di quelli effettuati dai restanti motori di ricerca che hanno visitato il nostro sito web.

Si può quindi dedurre che lo spider di questo motore di ricerca compie mediamente un numero di accessi nettamente superiore rispetto agli altri di cui si è cercato di analizzare il comportamento. Questo aspetto è molto importante, come già accennato, per quei siti web che necessitano di cambiare i propri contenuti velocemente e far in modo che su queste modifiche avvenga l'indicizzazione del sito o della pagina web. Nel nostro caso naturalmente, durante cinque settimane si è verificata una sola modifica per sito web; bisogna tenere conto che comunque i siti web attuali richiedono aggiornamenti sempre più costanti.

Si pensa ad esempio a portali turistici che interessano particolari zone d'interesse: se in un determinato periodo dell'anno si tengono manifestazioni o festival, è desiderabile che il sito web sia indicizzato per particolari parole chiave legate a quel contesto. In questi casi quindi è molto importante poter contare su un motore di ricerca che permetta di migliorare il posizionamento del proprio sito il più velocemente possibile, possibilmente non superando la settimana d'attesa. Se a un motore di ricerca occorre più tempo per capire che sono state apportate modifiche al sito web e decidere se indicizzarlo o meno, questo può rivelarsi dannoso per gli scopi prefissati da chiunque voglia indi-

cizzare un sito web, oppure può rendere necessario effettuare le modifiche richieste con un largo anticipo rispetto a quando queste possono rivelarsi realmente utili.

Per quanto riguarda motori di ricerca come Ask.com e Live Search, osservando i dati statistici relativi alle visite di robot e spider, si può notare come questi non abbiano visitato tutti i siti web creati, ma soltanto due a testa. Inoltre, tra i due, soltanto Ask.com ha indicizzato un sito web per un totale di cinque parole chiave. Poiché è stato appurato che, comunque, gli spider di questi due motori di ricerca hanno visitato due siti su sei, non è possibile affermare che la mancata visita negli altri quattro domini sia da attribuire alla lentezza del lavoro degli spider. Piuttosto, si può ipotizzare che questa dipenda dalle tecniche di indicizzazione utilizzate che, benché si siano rivelate utili per Google, possono non portare agli stessi risultati per gli altri motori di ricerca.

Inoltre, anche i siti web che hanno ricevuto la visita degli spider, non hanno ottenuto un miglioramento della loro posizione all'interno della SERP di questi motori di ricerca. L'unico sito web indicizzato da Ask.com è stato quello in cui sono stati utilizzati i BackLink come tecnica d'indicizzazione.

Possiamo quindi ipotizzare che l'uso di questa tecnica sia apprezzata da questo motore di ricerca, ma che abbinata alle altre sperimentate non porti al medesimo risultato poiché per il sito [www.comeindicizzareunsitoweb.it](http://www.comeindicizzareunsitoweb.it), nonostante sia stato visitato dallo spider Inktomi Slurp di Ask.com, non sono stati registrati miglioramenti nel suo posizionamento all'interno del motore di ricerca.

Per quanto riguarda invece Live Search, possiamo dedurre che le tecniche utilizzate non siano adatte per l'indicizzazione all'interno di questo motore di ricerca poiché i siti web che sono stati visitati da MSNBot sono [www.lindicizzazione.it](http://www.lindicizzazione.it), sito web a cui non sono state applicate tecniche d'indicizzazione, e [www.ilrankingweb.it](http://www.ilrankingweb.it), ottimizzato tramite i BackLink.

## **EFFETTI DELLE TECNICHE DI OTTIMIZZAZIONE**

Tra le tecniche di ottimizzazione utilizzate, si può osservare che quelle che hanno apportato dei netti miglioramenti nel posizionamento dei siti web all'interno dei motori di ricerca sono essenzialmente due. In particolare, possiamo dire che la tecnica più efficace che è stata sperimentata riguarda i BackLink; questo è deducibile dal fatto che il sito [www.ilrankingweb.it](http://www.ilrankingweb.it) è quello che ha ricevuto il maggior numero di visite da parte degli spider. Questo dominio infatti è l'unico che ha ricevuto visite da quasi tutti gli spider dei motori di ricerca presi in considerazione (tranne Yahoo!) ed è inoltre l'unico sito web ad essere stato indicizzato in Ask.com.

Efficace si è rivelata anche la tecnica dei meta tag, che ha portato all'indicizzazione di 9 parole chiave in Google, due in più rispetto a [ilrankingweb.it](http://ilrankingweb.it), anche se il posizionamento per le parole chiave indicizzate è più basso rispetto a quest'ultimo.

Il sito web indicizzato tramite l'utilizzo del tag HTML invece è stato ottimizzato soltanto per tre parole chiave all'interno del motore di ricerca Google, guadagnando comunque posizioni molto basse, ad eccezione della parola chiave "meta pragma", che inspiegabilmente è salita al secondo posto.

Per quanto riguarda l'uso combinato delle diverse tecniche, si può dire che in Google queste hanno portato all'indicizzazione del maggior numero di parole chiave, anche se il posizionamento, in taluni casi, è peggiorato.

L'uso delle sitemap si è invece rivelato fallimentare per il miglioramento di posizioni del sito web; il dominio ottimizzato con questo metodo infatti non solo non è stato indicizzato, ma in alcuni casi è stata addirittura peggiorata la posizione che per le stesse parole chiave si era guadagnata sul sito web non ottimizzato.

Tuttavia, l'uso di sitemap si è rivelato utile nel contesto del motore di ricerca Google, soprattutto per far sì che questo indirizzi sul nostro sito web il suo spider ogni qualvolta venga effettuata una modifica alla pagina. È quindi una tecnica utile per aumentare le visite di GoogleBot al proprio sito web, ma di per sé non permette l'ottimizzazione del sito stesso. Possiamo in conclusione affermare che inserire una sitemap all'interno del sito, sia in formato HTML che in formato XML, è in ogni caso utile per il lavoro di un web master, purché questa sia affiancata da altre tecniche di ottimizzazione.

Un discorso a parte merita il motore di ricerca Yahoo!; questo infatti è l'unico che non ha effettuato alcuna visita a nessuno dei siti web creati. E' quindi ipotizzabile che le tecniche d'indicizzazione utilizzate non siano state efficaci per questo particolare motore di ricerca; in ogni caso va notato che al momento della sottomissione dei siti web a Yahoo!, il motore di ricerca specificava che un'indicazione gratuita del proprio sito web non garantiva il passaggio del proprio spider all'interno del sito stesso. Si può affermare che i metodi più efficaci e veloci per fare in modo che un sito web salga anche all'interno della SERP di questo motore di ricerca siano i prodotti messi a disposizione a pagamento.

Si può quindi concludere che, tra le tecniche analizzate e sperimentate, sicuramente l'uso di link e BackLink è quella che si è rivelata più efficace per gli scopi prefissati. Naturalmente l'esperimento effettuato è stato volutamente circoscritto in un periodo di tempo determinato, sia per stabilire il "tempo di reazione" dei motori di ricerca alla pubblicazione dei siti web, sia per testare, nel caso di avvenuta indicizzazione, quanto tempo occorresse alle tecniche utilizzate per ottimizzare il posizionamento dei siti web all'interno di una SERP.

## WEBLIOGRAFIA E SITOGRAFIA

### Webliografia di riferimento

- 1 Rand Fishkin of SEOmoz.org, *Beginner's Guide to SEO*, Guida presente su SEOmoz.org
- 2 Taher H. Haveliwala (2003), *Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search*, Stanford University, Conferenza WWW2002
- 3 Taher H. Haveliwala (2002) , *Topic-Sensitive PageRank*, Stanford University, Computer Science Department, Conferenza WWW2002
- 4 Sergey Brin and Lawrence Page (1998), *The Anatomy of a Large-Scale Hypertextual Web Search Engine*, Computer Science Department, Stanford University
- 5 Junghoo Cho, Hector Garcia-Molina, Lawrence Page (1998), *Efficient Crawling Through URL Ordering*, Department of Computer Science Stanford University, Rivista "Computer Networks and ISDN Systems", volume 30

### Sitografia di riferimento

- 6 [www.seoutilility.com](http://www.seoutilility.com)
- 7 [www.giorgiotave.com](http://www.giorgiotave.com)
- 8 [www.motoricerca.info](http://www.motoricerca.info)
- 9 [www.seomoz.org](http://www.seomoz.org)
- 10 [www.googlerank.com](http://www.googlerank.com)