

Prefazione

La tesi analizza il problema del posizionamento di un sito all'interno dei motori di ricerca, nel caso specifico viene considerato il sito della facoltà di Scienze Statistiche.

Lo scopo di questa tesi è di capire se le pagine web siano create correttamente, e se la loro disposizione all'interno del sito sia adeguata alle richieste delle varie tipologie d'utenti.

Il sito deve essere creato in maniera tale che i visitatori che non lo conoscono, siano incentivati a farlo, senza fare una politica di posizionamento lesiva.

Il ricorso alle indagini statistiche, invece, contribuisce a individuare quali sono le pagine più linkate, al fine di porre in risalto quelle meno visibili e correggere eventuali errori di struttura logica del sito.

INDICE

Capitolo 1: Introduzione	5
1.1 Come realizzare una campagna di posizionamento	7
1.2 Strategia per un posizionamento ottimale	10
1.3 Webpromotion	13
1.4 Considerazioni	15
Capitolo 2: Posizionamento	16
2.1 Differenza tra indicizzazione e posizionamento	16
2.2 Link, testo e popolarità	18
2.3 Ranking	23
2.4 Keywords	25
2.5 Considerazioni	28
Capitolo 3: Statistiche	30
3.1 Activity Statistics	33
3.2 Access Statistics	35
3.3 Browsers	37
3.4 Errors	39
3.5 Considerazioni	39
Capitolo 4: Conclusioni	41
Capitolo 5: Bibliografia	46
5.1 Libri e articoli	46
5.2 Dati e relative tabelle	47
5.3 Glossario	50

CAPITOLO 1: INTRODUZIONE

Prima di addentrarci nel caso specifico preso in esame, è utile capire l'importanza del posizionamento.

Il posizionamento è l'insieme d'operazioni che hanno l'obiettivo di far comparire un sito in testa ai risultati dei motori di ricerca, utilizzando quelle parole chiave (o, meglio, combinazioni di termini) che riguardano l'attività del sito stesso. Anche se non è veramente importante apparire obbligatoriamente ai primi posti, perché l'obiettivo principale è quello di attirare il traffico di qualità.

Il 70% degli utenti non consulta più di tre pagine di risultati dopo una ricerca sui motori. È quindi fondamentale che il link al proprio sito si collochi tra i primi risultati e in più motori e directory possibili per avere un'adeguata visibilità e "popolarità".

A tal proposito, ho fatto una ricerca sui principali motori e directory per vedere se realmente il sito della facoltà ha una visibilità soddisfacente.

MOTORI DI RICERCA

PAROLE DIGITATE	GOOGLE	MSN	ALTAVISTA	EXCITE	LYCOS
facoltà statistica	1° posto	1°	1°	4°	1°
Statistica facoltà	9° posto	1°	1°	4°	1°
lauree statistica	27° posto	8°	9°	9°	6°
master statistica	14°posto	7°	27°(Al)	Nc	Nc
corsi statistica	Nc	11°	8°(B)	Nc	7°(B)
faculty statistical	18°posto	1°	1°	1°	1°

Nc= non compare nelle prime 3 pagine; B= compare il sito della biblioteca; Al= compare sito Aledis

WEB DIRECTORY

PAROLE DIGITATE	FREEONLINE	VIRGILIO	SUPEREVA	YAHOO	DMOZ
facoltà statistica	3° posto	3°	15°	1°	1°
statistica facoltà	Nc	9°	3°	1°	1°
lauree statistica	27°posto	Nc	5°	9°	8°
master statistica	26°posto	8°	24°(AI)	21°(AI)	Nc
corsi statistica	18°posto	Nc	9°(B)	8°(B)	Nc
faculty statistical	9°posto	9°	1°	1°	Nc

Nc= non compare nelle prime 3 pagine; B= compare il sito della biblioteca; AI= compare sito Aledis

I numeri all'interno delle tabelle fanno riferimento al posto in cui si trova il sito della Facoltà sui motori o directory (posizionamento). In base ai dati raccolti si denota che nella maggior parte dei casi compare l'homepage, ma digitando le parole:

- corsi statistica viene visualizzata la pagina della biblioteca;
- master statistica compare il sito Aledis (Associazione Laureati e Diplomati in Statistica).

Nelle Web Directory è più frequente il valore nullo, rispetto ai motori di ricerca, digitando le medesime parole.

Quello che ci conforta è che sul principale motore di ricerca (Google), il sito è presente nelle prime 3 pagine.

Da questa analisi può sorgere un dubbio: "è più utile essere posizionati su un MDR o su una Directory?"

Innanzitutto è d'obbligo capire il loro scopo:

- i motori, scandagliano continuamente l'intero www e includono nel proprio archivio di indirizzi tutti i siti Web trovati a prescindere dalla qualità dei loro contenuti;

- una directory contiene una raccolta di indirizzi di siti web che sono stati espressamente selezionati da personale umano.

In generale si può dire che è preferibile essere “popolari”, cioè ben visibili in entrambi, perché ci sono varie tipologie di utenti. Quest’ultimi, quando hanno ben chiaro l’elemento da cercare, si avvalgono dei motori di ricerca, mentre quando vogliono trovare le varie alternative preferiscono le Web directory.

Ad esempio, se un visitatore cerca l’indirizzo internet della Facoltà di Scienze Statistiche di Padova, utilizza un motore di ricerca, mentre se l’utente vuole trovare la lista delle maggiori Facoltà di Scienze Statistiche nazionali, utilizzerà una Web Directory.

1.1 Come realizzare una campagna di posizionamento

La campagna di posizionamento prevede fasi precise:

Definizione delle parole chiave e dei motori

Il primo passo consiste nella definizione delle parole chiave per le quali verranno ottimizzate le pagine. Vanno definite in piena collaborazione tra il consulente e il gestore del sito perché saranno il "biglietto da visita" del sito stesso. In generale è opportuno che:

- non siano più di 5 o 6 per ciascuno dei macroargomenti trattati dal sito per non disperdere gli sforzi in più direzioni.

- non siano troppo generiche al fine sia di trapiandare un pubblico specifico sia di inserirsi in un ambiente meno "competitivo" e che non richieda troppa manutenzione. E' comunque da rilevare che, poiché la lingua più diffusa in Internet è l'Inglese, i vocaboli italiani raramente danno problemi di posizionamento.
- siano accompagnate da una o due parole chiave secondarie, che normalmente sono una frasetta che comprende la keyword principale.

In questa sede vengono anche decisi i motori di ricerca per cui ottimizzare. Poiché il 90% del traffico diretto ai motori di ricerca è realizzato dai 10 motori principali, è del tutto inutile prendere in considerazione piccoli motori che generano poco o nessun traffico. Può essere utile, però, in alcuni casi particolari, trapiandare alcuni motori di argomento specialistico o di ambito regionale che abbiano anche una discreta popolarità.

Realizzazione delle pagine

Il consulente, sulla scorta delle informazioni fornite, realizza una pagina di presentazione del sito da utilizzare come base per la costruzione dei gateway. Approvato il testo, questo viene inserito nelle pagine gateway con modifiche di forma e non di contenuto che sono necessarie ai fini dell'ottimizzazione. Il numero delle pagine da realizzare è pari a: (n° motori scelti) x (n° di parole chiave scelte). Le pagine vengono quindi sottoposte all'approvazione del gestore del sito e caricate sul web server in modo da essere accessibili tramite Internet, ed il consulente le sottopone ai motori di ricerca.

Reporting

Con cadenza variabile vengono sottoposti al gestore del sito dei report che consentono di seguire l'andamento del posizionamento sui motori scelti. Qualora qualcuna delle pagine non dovesse dare i risultati attesi il consulente procede agli aggiustamenti necessari. I motori impiegano un tempo variabile tra le 2 settimane e i 3 mesi per indicizzare la pagina, per cui questa fase, compreso qualche piccolo aggiustamento che si renda necessario, si può considerare conclusa al massimo 3 mesi dopo le prime sottoposizioni.

Manutenzione

Il posizionamento di una pagina su di un motore non è costante nel tempo ma può subire delle variazioni. Ciò è dovuto alle piccole modifiche al software che ogni motore applica per mantenersi aggiornato e, soprattutto, alla sottoposizione di nuove pagine che ricevono una classificazione migliore. Se il gestore lo ritiene opportuno, è possibile procedere, con una frequenza variabile tra il settimanale ed il bimestrale, al controllo del posizionamento e ad effettuare le correzioni necessarie. Il gestore può comunque tenere sotto controllo le statistiche relative agli accessi al proprio sito e valutarne l'andamento nel tempo.

1.2 Strategia per un posizionamento ottimale

Il fatto di focalizzarsi su una nicchia ben precisa aiuta molto: avere un prodotto o un servizio unico con pochi concorrenti presenti è quasi una garanzia per un buon piazzamento; nel nostro caso la concorrenza è data dalle altre Facoltà di Scienze Statistiche presenti in Italia (eventualmente Europa). L'insieme dei siti relativi ad argomenti omogenei presenti nell'indice di un motore di ricerca rappresenta il "core". I motori fanno l'analisi di tutte le pagine di ciascun "core" e calcolano un Page vector, in pratica danno un "voto" a ciascuna pagina presente nel loro indice e relativa ad un certo argomento. Da notare l'importanza della Differenziazione: se siete in una categoria competitiva dovete trovare il modo di differenziarvi altrimenti correte il rischio di rimanere confusi nella massa dei concorrenti.

A proposito di pagine, la loro specializzazione aumenta la rilevanza delle parole chiave di ricerca perché testo, link e META tags sono completamente dedicati ad uno o pochi specifici prodotti invece di disperdersi in varie direzioni. La facoltà rende evidente i servizi che offre, in conformità a ciò che si sceglie si hanno argomenti specifici.

Per quanto riguarda l'homepage, la maggior parte dei MR non indicizza pagine che sono senza testo (ma, ad esempio, solo grafici). È utile far notare la differenza tra una pagina introduttiva e una "doorway pages": quest'ultime sono costruite utilizzando un numero di parole più piccolo possibile, spesso senza dire nulla del contenuto del sito a cui fanno riferimento. Hanno semplicemente una singola immagine e un link al resto del sito. Le pagine introduttive, invece, hanno un maggior

contenuto delle “doorway pages” e sono utilizzate principalmente per due scopi:

- 1) fornire ai motori di ricerca una ragionevole introduzione sul contenuto che i visitatori possono trovare visitando le pagine;
- 2) aiutare i siti con un buon contenuto di essere posizionati prima di quelli che contenuto non lo hanno.

Una particolare attenzione va data alle “Pagine Orfane”: se un MR ha incluso nell’indice una pagina che e’ stata successivamente cancellata, il visitatore che clicca il link si ritrova con un avvertimento che “la pagina richiesta non e’ rintracciabile”.

Come capite bene, meglio lasciare attive tutte le pagine registrate, al massimo cambiate il contenuto se dovete, ma assolutamente non cancellatele dal sito. Una pagina orfana, mette in cattiva luce il sito. Il sito di Statistica non presenta le suddette pagine.

Un problema relativo alle pagine, può derivare dal Tag <Redirect>. È un tag HTML che specifica su che pagina ridirigervi e dopo quanti secondi. La maggior parte dei MR non indicizzano pagine che includono il tag, infatti il sito di Scienze Statistica non lo utilizza.

Nella scelta delle parole/frasi di ricerca, è consigliabile provarle prima sui principali motori di ricerca e vedere se i siti che compaiono sono inerenti al servizio che offrite.

È utile oltre alle parole chiave, analizzate la “fonte” (l’HTML) dei siti meglio piazzati per ciascun MR. Il codice Html ci consente di vedere sia la densità delle keywords, sia particolari che rendono il sito in vetta alla classifica.

Si nota che analizzando il sito della Facoltà, nel meta tag di descrizione del contenuto c'è scritto: "Facoltà di Scienze Statistiche dell'Università di Padova", che chiarisce a pieno il contenuto della pagina.

E' bene far notare che "infarcire" il sito con parole di ricerca che non hanno a che fare con il vostro sito, non serve a niente. Il consiglio è di evitare trucchi, perché tanto non funzionano. Come tutte le cose sono anche, le parole di ricerca sono soggette alla legge della domanda e dell'offerta. Alcune parole sono molto competitive perché molto usate e sarà difficile raggiungere un buon piazzamento.

1.3 Webpromotion

In questo paragrafo analizzeremo le fondamentali regole che ci permetteranno di fare una politica non lesiva nei confronti del sito che vogliamo creare.

Innanzitutto dobbiamo scrivere nel title parole pertinenti (non utilizzare più di 60 caratteri) cominciando con una dall'iniziale A e possibilmente in minuscolo, maiuscolo, e con le iniziali maiuscole. Se consideriamo il tag di descrizione deve essere non superiore a 150 caratteri.

Title: "Facoltà di Scienze Statistiche – Università di Padova"

Descrizione: "Facoltà di Scienze Statistiche dell'Università di Padova"

Bisogna fare un uso strategico delle keyword, meglio se a coppie (non più di 150 caratteri). Ad esempio se riguarda Facoltà Statistica Padova, nella prima linea mettete Facoltà poi Facoltà Statistica infine nella terza lista Facoltà Statistica Padova:

- a. la parola a cui volete dare più rilievo (Statistica) inserirla nel codice HTML nei luoghi più disparati oltre i META tag, nel paragraph heading `<h2> </h2>`, nel source tag all'interno del campo ALT, scrivetele in minuscolo, le iniziali maiuscole e il resto minuscolo o interamente maiuscole, non ripetendo la stessa keyword all'interno di una sezione. Non usate parole d'uso comune;
- b. ripetere le parole chiavi anche nel testo della pagina nella parte superiore (e più a sinistra in caso di tabelle) possibile, contano maggiormente i primi 100-1000-2000 caratteri e anche la prima riga trovata;

- c. utilizzare le parole chiavi anche nei nomi di directory e files della struttura del sito.

Per non creare problemi di ricerca si cerchi di fare un sito il meno profondo possibile, avere dei contenuti facilmente reperibili e aggiornati. È utile utilizzare usare dei piccoli titoli occhiello che attirino il visitatore e nello stesso tempo lo spider.

Nell'homepage della Facoltà è sempre visibile un link con le principali directory presenti nel sito. Per far ciò è stata utilizzata una tabella, o non i frame che potrebbero causare problemi nel caricamento della pagina web. Oltre ai frame, non usate:

- Intestazioni grafiche;
- Lead Page, pagine che contengono tutte le malizie possibili;
- Auto Forwarding Page, pagine che automaticamente inviano l'utente ad un'altra page.

Infine e non meno importante rendere il sito più visibile facendosi linkare attraverso una politica di cross promotion; come il sito effettua con le scuole superiori.

1.4 Considerazioni

A prima vista può apparire paradossale, ma bisogna concepire il sito per i propri utenti e non per i motori di ricerca. L'ingrediente principale sul quale applicare le tecniche di posizionamento sui motori di ricerca, è proprio il contenuto: esso deve risultare interessante, articolato e deve offrire un reale servizio informativo.

Oltretutto quanto più un sito offre contenuti di rilievo, tanto più avrà la possibilità di diventare "popolare" e ottenere link da siti esterni.

La filosofia di Google, il più importante motore di ricerca al mondo: il numero di link ottenuti (link popularity) esprime il livello d'importanza e di popolarità di un sito e costituisce uno dei principali fattori usati per determinare il Page Rank delle pagine. Più è elevato il Page Rank, maggiori possibilità ha un sito di comparire tra i primi risultati.

Per queste ragioni, un'efficace campagna di web marketing e di posizionamento sui motori, deve coinvolgere figure professionali in grado di coprire più livelli di esperienza, tra i quali:

- Programmatori Html con conoscenze specifiche sui search engines;
- Content Manager (per il concepimento dei contenuti del sito);
- Copywriter pubblicitari (per la definizione di messaggi mirati ed efficaci);
- Conversion management (per realizzare specifici contenuti);
- Esperti di analisi del traffico web

CAPITOLO 2: POSIZIONAMENTO

In questo capitolo analizzeremo più in profondità gli argomenti espressi nell'introduzione.

Molti credono che tutto finisca una volta che il designers ha realizzato la parte grafica del sito, ma non è così perché realizzare un sito non vuol dire per nulla essere trovati. Per questo bisogna saper adattare i siti alle esigenze dei motori di ricerca, per poter effettuare l'indicizzazione con maggiore sicurezza. Per chiarire perché è stato citato il termine indicizzazione e non posizionamento, il paragrafo seguente spiega le loro differenze.

2.1 Differenza tra posizionamento e indicizzazione

Quando si naviga in Internet, spesso si sente “giocare” con i termini indicizzazione e posizionamento. Ciò dipende dal fatto che si tratta di parole che consentono di collocarsi ai primi posti nei Motori di Ricerca.

Il rispetto di potenziali clienti (l'etica) e forse anche la conoscenza della lingua italiana imporrebbe un comportamento completamente differente giacché i due differenti servizi Web Marketing possono avere costi molto diversi l'uno dall'altro.

Partiamo dai concetti elementari, le definizioni per poi passare alle analisi.

L'indicizzazione consiste nell'inserimento dell'URL di un sito (<http://www.nomesito.com/>) nelle banche dati dei Motori di Ricerca e delle Directory.

Il posizionamento consiste nell'effettuare questa prima fase, mettendo in atto poi una serie di strategie per far arrivare e mantenere tra i primi venti o trenta posti il proprio sito Web.

I Motori di Ricerca cambiano regolarmente (ogni mese circa) le regole in base alle quali stabiliscono le posizioni dei siti nelle proprie banche dati. Con la sola indicizzazione non si può mantenere il proprio sito tra i primi venti, ma si può solo sperare che ciò avvenga. Con il posizionamento queste garanzie aumentano considerevolmente.

Si sta diffondendo la tendenza ad offrire una sorta d'indicizzazione di base ed una più completa. La prima comporta l'inserimento del proprio sito unicamente nei database dei principali Motori di Ricerca e Directory. Quella più completa, invece, consente l'inserimento del proprio sito in un elevato numero di data base di Motori e Directory (nazionali ed internazionali), persino alcune centinaia.

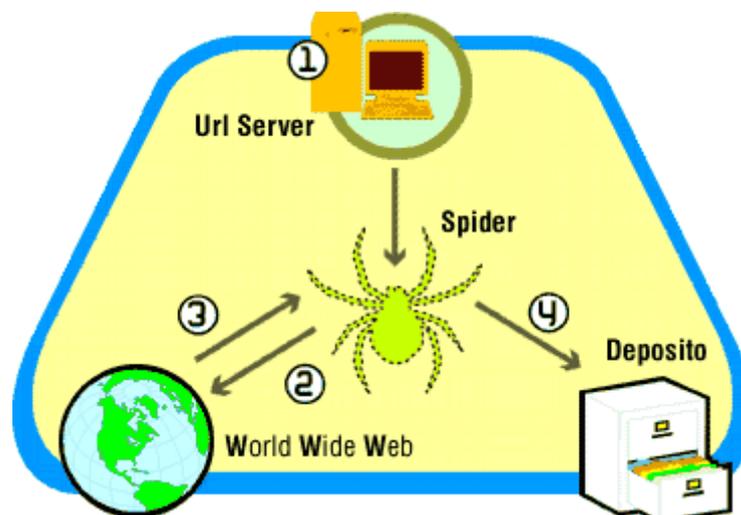
Registrarsi in alcune centinaia di Motori, peraltro, c'espone ad un forte rischio "spam" ,siccome alcuni Motori sono solo in realtà dei mezzi per raccogliere dati personali (i vostri dati...), da rivendere al miglior offerente.

Bisogna quindi stare attenti, e usare il giusto compromesso tra un'indicizzazione completa, inserendo il proprio sito su motori conosciuti per la loro affidabilità, e non su quelli di scarso utilizzo.

2.2 Link, testo e popolarità

A proposito di motori di ricerca e directories, online n'esistono moltissimi, ognuno dei quali enfatizza in modo preponderante uno dei tre fattori elencati nel titolo qui sopra. Ciò dipende dall'algoritmo che sta alla base degli spiders (robots), o delle norme seguite da chi si occupa dell'inserimento dei siti nelle Directories.

Di seguito riportiamo una figura che spiega la funzione degli spiders:



Ipoteticamente, il sito che ottiene il miglior ranking è quello che ottiene il miglior risultato possibile per ogni aspetto elencato sopra (testo, link e popolarità) senza sfociare nello spamdexing, rischio di chi si occupa d'ottimizzazione di pagine web. Per quanto riguarda il ranking lo tratteremo nel prossimo paragrafo.

Per capire meglio il vero valore dei tre fattori, li analizzeremo singolarmente:

LINK

Molto probabilmente se non ci fossero, non ci sarebbe la Rete, poiché è grazie ad essi che si può passare da una pagina all'altra. Vorrei subito far notare una distinzione fondamentale: link e banner. Il link è una parola o frase che, cliccata collocandovi sopra il puntatore del mouse, permette di raggiungere una pagina web di destinazione. Con il banner il principio è il medesimo solo che al posto del testo si usa un'immagine di dimensioni variabili, animata e non.

Se dobbiamo considerare l'ottimizzazione e di conseguenza il posizionamento, devo far notare che i link sono da preferire giacché sono seguiti dagli spiders mentre i banners no, al massimo sono seguiti dai visitatori interessati a raggiungere la pagina di destinazione.

Nel momento in cui s'impone un link bisogna scegliere in modo efficace le parole chiave che lo costituiranno, il cosiddetto microcontent che si differenzia dal resto del testo definito macrocontent. Il microcontent andrà scelto in funzione della pagina di destinazione, ossia dovrà esserne una sorta di riassunto.

I link si possono sfruttare per aumentare il traffico del proprio sito web, anche se il miglior modo per ottenere questo risultato è un buon posizionamento nei Motori di Ricerca, oltre che un inserimento nelle principali Directories.

Tra le tecniche principali ricordiamo lo scambio link tra siti preferibilmente con argomento affine. A linka B, B linka A ed il gioco è fatto (cross linking).

Alla base di queste metodiche e di altre anche più complesse ci sono due scopi principali: l'aumento del ranking ed anche quello del traffico. Quale dei due fattori sia più importante credo sia difficile dirlo, molto dipende dallo scopo che ci si prefigge con il proprio progetto Web in un determinato momento: aumentare il traffico per aumentare i propri potenziali clienti o, magari, migliorare il proprio ranking alla ricerca di traffico che manca.

Una corretta analisi di partenza è in grado di determinare quale sia l'obiettivo più importante da raggiungere per il successo del proprio progetto Web.

Analizzando il sito, si riscontra che esso è suddiviso in due blocchi (è stata utilizzata una tabella, niente frame): il primo è sostituito da un menu sempre presente durante la navigazione; il secondo illustra i contenuti in base alle scelte che ha effettuato l'utente.

Nel menu sono presenti vari link divisi per aree:

- Conoscere la facoltà: corsi di studio, progetti per gli studenti, news ed eventi;
- Servizi per gli studenti: comunicazioni, programmi e appelli d'esame, lauree, modulistica e forum. Per accedere ai suddetti servizi bisogna essere registrati al sito;
- Servizi per docenti (IRIDE).

Nell'ambito dei link abbiamo dei collegamenti utili verso il sito della Biblioteca, il programma Erasmus e l'Associazione Aledis.

Per quanto riguarda i link esterni, la facoltà ti permette di accedere allo sportello FSE (fondo sociale europeo) e a tutta l'Università agli studi di Padova.

TESTO

Oltre a link e popolarità, il testo è un elemento fondamentale poiché ormai i Motori di Ricerca tralasciano sempre più spesso i meta-tags con i dati impostati da noi, per badare ai contenuti nella maggior parte dei casi scarni.

La maggior parte dei designers bada all'immagine senza considerare che un sito deve essere trovato e, perché ciò accada, devono verificarsi delle condizioni importanti una delle quali è la ricchezza di contenuti in relazione alle parole/frasi chiave.

Se prima era fondamentale prestare la massima attenzione nella collocazione/scelta delle parole e delle frasi chiave per i meta-tags, ora bisogna dedicare il massimo dell'attenzione ai contenuti, la parte visibile di un sito, diversa dai meta-tags non visibili online dai navigatori.

Partiamo dai Motori e dai fattori ai quali fanno riferimento per il ranking dei siti web. In relazione ai testi considerate che vista la mole di pagine quotidianamente indicizzate, si preferisce immagazzinare solo la parte iniziale d'ogni pagina. Pertanto, la scelta e la collocazione delle parole/frasi chiave in inizio di pagina sarà fondamentale. Altri fattori rilevanti che consentono di attribuire maggior "peso" ad una parola/frase sono l'utilizzo del **grassetto**, del sottolineato e del *corsivo*.

Si possono utilizzare gli alt-tag per inserire altro testo, in alcuni casi si consiglia di riferirsi alla stessa description e certamente alle keywords elencate nei meta-tags. Approfitto della considerazione per fare notare che è molto meglio non inserire keywords nei metatags se poi non si utilizzano nel proprio testo.

Altro aspetto da curare è la keyword proximity, ossia la vicinanza tra parole chiave. Inoltre, la presenza di link e la scelta del loro testo influenza in modo globale il ranking del sito. Per evidenziare delle parole e delle frasi di rilievo si possono utilizzare anche H1 e tutti gli altri.

Probabilmente il sito del futuro sarà esente da trucchi e trucchetti vari, visto che i Motori sono sempre più accorti da questo punto di vista. Chi saprà conoscere giorno dopo giorno i Motori uscirà vincente, invece, chi avrà cercato di migliorare il proprio posizionamento online solo con trucchi vari ne uscirà perdente.

Nell'homepage del sito e nelle successive pagine, il testo è accompagnato da piccole immagini, che mettono meglio in risalto il contenuto, ad esempio del servizio scelto. In particolare il titolo è in grassetto per attirare l'attenzione da parte dell'utente. Sotto di esso, appare una breve didascalia con affianco una figura.

POPOLARITA'

La popolarità di un sito è lo strumento fondamentale per il successo dello stesso, poiché è dall'impostazione di tutti i suoi fattori che ne dipende il successo. Due sottoelementi definiscono questo importante fattore: link popularity e click popularity.

La link popularity indica il numero di siti diretti verso il sito in esame, mentre la click popularity indica quanti click riceve il sito e per quanto tempo rimane ogni visitatore.

Questi fattori combinati insieme determinano il ranking di un sito. Dobbiamo considerare che il diverso algoritmo per ogni Motore di

Ricerca, a parità di link popularity e click popularity da' come risultato diversi ranking.

Lo scambio link è un fattore molto utile per aumentare la popolarità di un sito ma non è uno dei fondamentali. Potremmo definirlo complementare, adatto se si considera un lungo termine. Non si confonda lo scambio link con lo scambio banner, poiché i Motori di Ricerca non seguirebbero i banner.

Le tecniche che ho evidenziato riguardano la popolarità di un sito, in relazione ai Motori di Ricerca e non ai visitatori. Nel caso in cui si considerano entrambi, allora può essere interessante anche uno scambio banner preferibilmente con traffico targettizzato.

2.3 Ranking

Nel paragrafo precedente lo abbiamo citato varie volte nell'ambito dei tre fattori (link, testo e popolarità).

È il metodo con cui un motore dispone e dall' algoritmo dipende il successo di un motore di ricerca. Nel passato un motore non utilizzava degli algoritmi particolari per pubblicare una lista di risultati, ma assegnava direttamente un valore d'importanza alle pagine restituite in base al numero di ricorrenze, in ogni documento, dei termini della query. Naturalmente, in un ambiente particolare come il web, attraverso questo metodo era facile apparire alle prime posizioni: bastava inserire un numero elevato, ed a volte esagerato, di parole chiave dentro una

pagina e il gioco era fatto. Questo sistema, però, è durato poco ed i motori di ricerca hanno cominciato a bannare (eliminare dai loro db) e considerare spamming (in pratica, tecniche fraudolente utilizzate per eludere il motore di ricerca e fare apparire, in ogni caso, il proprio sito nei primi posti) alcuni metodi di costruzione delle pagine.

Inizialmente i motori assegnavano importanza alle pagine in base al numero di parole che l'utente ricercava. Per ovviare a questo sistema, una tecnica moderna di ranking è quella della link popularity, in base alla quale un valore d'importanza era dato dal numero di link che una pagina riceveva dalle altre, sia interne sia esterne al sito.

In origine, alla base di tutto c'era il problema/obiettivo di riuscire ad usare un parametro umano per determinare quanto fosse rilevante una pagina sulle altre, rispetto ad un certo termine. “Ma come fare questo senza dover coinvolgere (e quindi pagare) un gran numero di persone e studiare un sistema valido per consentire loro di dare valutazioni oggettive?” L'unico modo era creare un sistema matematico incentrato sulla valutazione, spesso non espressamente riconosciuta, di un sito da parte di altri responsabili, amministratori e creatori di website.

In pratica, il sistema doveva poter rispondere alle domande: cosa pensano gli altri del mio sito? Quanto lo valutano pertinente e interessante rispetto ad un dato argomento?

Google sembra essere stato un pioniere nell'utilizzo dei collegamenti ipertestuali, per la creazione di un sistema di valutazione dell'importanza di un website. Grazie alla tecnologia chiamata PageRank, Google riesce oggi a fornire risultati molto pertinenti, stabilendo in modo del tutto automatizzato una misurazione oggettiva dell'importanza di ogni singola pagina web presente nel suo database.

2.4 Keywords

Attualmente la maggior parte dei Motori di Ricerca bada molto al contenuto presente tra i tag body. Questa regola non vale per tutti, è quindi buona norma realizzare i tag per le keywords.

Keyword sta per parola chiave ed indica una parola attorno alla quale viene costruito tutto (o parte...) del contenuto di una pagina Web. Analizzando qualunque sito Web, pagina per pagina, noterete che nella parte del codice troverete sempre un tag dedicato alle keywords, non esiste una sola pagina che faccia da riferimento per le parole chiave.

Altro fattore da considerare è che per loro natura le parole chiave devono essere un riferimento per i Motori di Ricerca che si basano su di esse. Differenziando le keywords per pagina si fa in modo tale che ogni pagina possa essere trovata singolarmente.

Se ammettiamo che le keywords siano uguali in tutte le pagine allora ammettiamo che tutte le pagine hanno lo stesso contenuto, visto che parole e frasi chiave devono essere lo specchio della pagina in cui si trovano. Invece, con parole chiave diverse per pagine differenti si ammette che le pagine hanno contenuto diverso.

All'inizio sembra solo una spesa onerosa quella di scegliere le parole chiave, ma in realtà è un investimento nel proprio business online.

Di seguito analizzeremo come scegliere le Keywords e dove posizionarle al meglio.

Scelta delle parole o frasi chiave

Tutto dipende dal prodotto con cui si ha a che fare e dalla concorrenza che esso ha con i suoi competitors (concorrenti).

Dal punto di vista prettamente teorico sappiate che le parole o, ancor meglio, le frasi chiave dovranno cercare di essere rappresentative del prodotto stesso evitando inutili e facili generalizzazioni. Nel nostro caso, i competitors sono le Università, nello specifico le Facoltà di Scienze Statistiche presenti nel territorio italiano. Importante sarà riuscire ad attirare l'attenzione dei cosiddetti "clienti", offrendo servizi che altre Facoltà non hanno.

Collaborano proprio per questo il proprietario del sito e consulente web nel momento cruciale della scelta delle parole e frasi chiave. In un primo momento il cliente dovrà scegliere da cinquanta a settanta parole o frasi chiave che ritiene più adatte. Successivamente questo materiale viene passato al consulente web che elimina parole o frasi che non inciderebbero in modo sostanziale sul posizionamento o che comunque non sono utili ai fini in considerazione.

La tendenza da parte degli optimizers meno esperti è quella di scegliere parole chiave più comuni, probabilmente perché danno un senso di sicurezza. Questo ragionamento è sbagliato, perché più una parola è comune e più è facile che ci sono moltissimi competitors. Soprattutto all'inizio, meglio definire parole e frasi chiave con pochi concorrenti garantendosi i primi venti o trenta posti nel ranking (vedi paragrafo precedente) dei principali Motori di Ricerca. Per il nostro sito, vengono utilizzate le seguenti parole chiave: Facoltà, Statistica, Università, Padova.

Mi preme far notare, che molte volte si tende ad abbandonarsi a campagne che allontanano dal proprio target anche a danno del lavoro fatto in precedenza. È utile rimanere in contatto con i potenziali clienti, pur restringendo il target.

Posizionamento delle parole chiave

Le parole chiave è consigliabile che siano collocate in diversi punti di una pagina Web per ottimizzarla al meglio, puntando ad un posizionamento quanto più efficiente possibile.

Con il passare del tempo, gli algoritmi che sono alla base di ogni Motore di Ricerca hanno cambiato i criteri di scelta dei siti Web, non basandosi più di tanto sulle keywords nella parte head ma sulle parole chiave nella parte Body del sito stesso.

Oramai si attribuisce sempre più importanza alla parte visibile dei siti (contenuti) a scapito di quella non visibile (parole chiave e description nella parte head).

La nostra pagina Web sarà valutata dai motori di Ricerca prendendo in considerazione sia la parte iniziale ed il titolo prescelto (*keyword prominence*), sia la presenza di parole chiave nei link presenti nel testo. A proposito di link, anche il numero di link è un fattore rilevante.

2.5 Considerazioni

Per concludere, si può dire che le pagine web, a differenza dei normali testi, possono avere dei link ad altre pagine, e questi link vengono interpretati dai motori di ricerca come una specie di voto da parte delle altre pagine, e quindi da parte degli altri siti e webmaster. Bisogna tener presente che un link ricevuto da un sito importante varrà sicuramente più di un link proviene da pagine meno importanti.

Per quanto riguarda le parole chiave bisogna prestare attenzione alla loro disposizione e numerosità all'interno del sito.

Nell'ambito del sito della Facoltà si nota che le keywords:

- sono al plurale: se utilizzi il plurale delle parole che hai scelto avrai più possibilità di essere trovati dopo una ricerca fatta su un motore di ricerca (ad esempio: corsi statistica, anziché corso);
- diversificate: una sola parola da solo attrarrà poche persone, infatti non tutti pensano allo stesso modo. Un semplice esempio (vedi tabella introduttiva) mette in risalto la differenza di posizionamento del sito digitando le stesse parole, invertendo l'ordine. Le parole in questione sono: facoltà e statistica;
- utilizzare forme diverse: non tutti faranno delle ricerche utilizzando i termini "facoltà statistica". Giusto o sbagliato c'è un alto numero di persone che potrebbero fare delle ricerche con i termini "master statistica", "corsi statistica". È bene far notare che molti utenti effettuano delle ricerche utilizzando delle frasi che meglio indicano il tema della loro ricerca;

- numerosità parole chiave: è opportuno mettere un numero non elevato di keywords per pagina. Ad esempio, esse possono perdere rilevanza e peso. Importante dire che più pagine ha il sito, più parole chiave puoi adoperare;
- lingue: in base al target di riferimento; nel nostro caso vengono adoperate 2 lingue (italiano, inglese).

Capitolo 3: Statistiche

Per la rilevazione dei dati di traffico si possono usare vari metodi. I tre più diffusi sono:

1. BBM, browser-based measurement (misurazione basata sul browser);
2. Site-centric measurement (misurazione centrata sul sito);
3. User-centric measurement (misurazione centrata sull'utente).

Quest'ultimo tipo di misurazione è quello che utilizzeremo per la nostra indagine statistica relativa al periodo (21 febbraio 2005 – 22 marzo 2005). Si tratta di un'indagine di mercato, basata su tre elementi tipici del settore:

- a) Un campione di popolazione significativo del tipo di utenza che si vuole misurare;
- b) Uno strumento per la rilevazione del comportamento degli individui che compongono il campione, in grado di monitorare attimo per attimo qualsiasi tentativo, riuscito o non riuscito, di navigazione all'interno del sito;
- c) Strumenti statistici studiati per effettuare proiezioni più o meno attendibili, che estendono alla totalità della popolazione i dati ricavati dall'analisi del campione.

Entrando nello specifico, il sito della Facoltà consente di accedere ai vari servizi attraverso un'apposita procedura di registrazione, che implica l'invio dei dati personali da parte dell'utente (profiling esplicito).

Ricordiamo che il profiling (profilazione) è l'attività per mezzo della quale una serie complessa di dati relativi ad utenti/clienti viene elaborata da

specifici programmi, per generare alla fine la segmentazione della propria utenza in gruppi omogenei di comportamento.

Per altri servizi, c'è il profiling implicito che consiste nel tracciamento del comportamento di utenti anonimi nel corso delle loro visite nel sito. Il tracciamento può avvenire sia tramite IP sia tramite cookie.

Gli indirizzi IP che identificano il cosiddetto “visitatore unico” sono 16930 in un mese; e prendendo come campione statistico i 50 indirizzi più utilizzati, si scopre che solo il 10% si collega all'interno della Facoltà (vedi tabella). Si deve tener presente che sia gli studenti iscritti a Statistica, sia i docenti possono collegarsi al sito ad esempio dalla propria abitazione.

Il login principale ci consente di individuare:

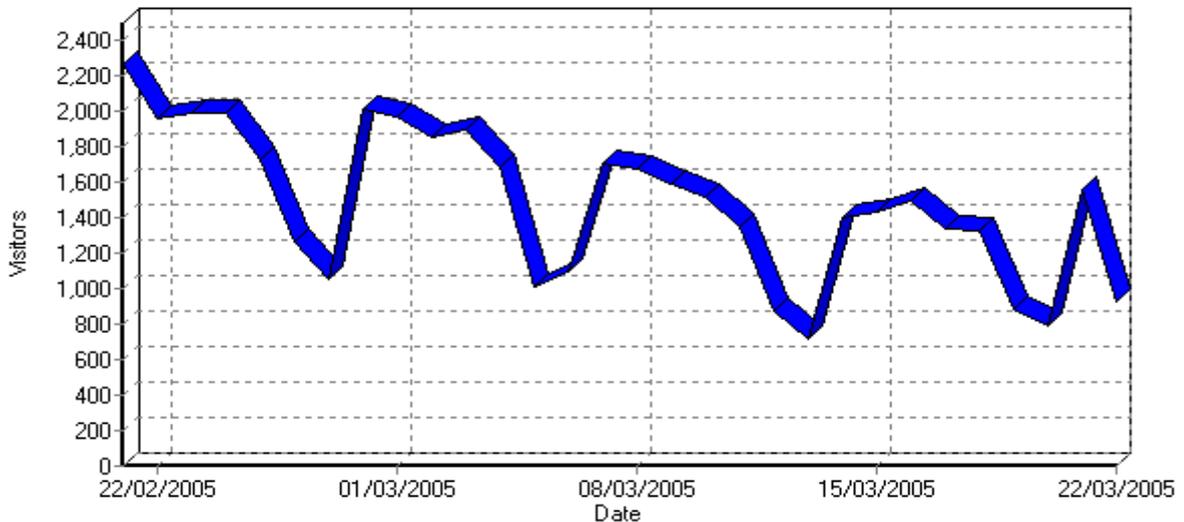
- nuovi visitatori: utenti mai registrati nei precedenti periodi di rilevazione statistica;
- visitatori perduti: l'elenco di utenti registrati, ma mancanti nelle rilevazioni del periodo corrente;
- visitatori di ritorno: utenti che hanno già visitato il sito in un precedente intervallo di tempo.

Indirizzi IP

	Host	Hits	Visitors
1	IANA (Internet Assigned Numbers Authority)	509	394
2	IANA	316	308
3	IANA	2,826	289
4	ESU-Padova	13,087	252
5	INFOSTRADA Internet	5,269	154
6	FASTWEB-VODAFONE-OMNITEL	4,150	147
7	FACOLTA' di STATISTICA	18,911	142
8	IANA(Internet Assigned Numbers Authority)	491	136
9	IANA	469	134
10	IANA	425	133
11	IANA	459	132
12	IANA	453	131
13	IANA	472	131
14	IANA	479	128
15	IANA	443	128
16	FACOLTA' di STATISTICA	30,413	128
17	TELECOM Italia Wireline Services	2,033	127
18	IANA (Internet Assigned Numbers Authority)	453	127
19	IANA	515	126
20	IANA	455	126
21	IANA	418	126
22	IANA (Internet Assigned Numbers Authority)	437	126
23	IANA	477	125
24	IANA	414	124
25	IANA	523	124
26	IANA	496	124
27	IANA	459	122
28	IANA	450	122
29	FACOLTA' di STATISTICA	5,377	122
30	REGIONE DEL VENETO	5,477	122
31	IANA (Internet Assigned Numbers Authority)	413	122
32	IANA	456	120
33	IANA	481	119
34	IANA	451	119
35	IANA	455	117
36	IANA	427	117
37	IANA	421	116
38	UNIVERSITA' di PADOVA	10,104	116
39	FACOLTA' di STATISTICA	30,120	115
40	IANA (Internet Assigned Numbers Authority)	450	114
41	IANA	452	113
42	FACOLTA' di STATISTICA	3,252	112
43	FACOLTA' di STATISTICA	5,748	111
44	UNIVERSITA' di PADOVA	14,392	110
45	IANA (Internet Assigned Numbers Authority)	461	109
46	IL TROVATORE s.r.l.	315	109
47	UNIVERSITA' di PADOVA	6,222	108
48	UNIVERSITA' di PADOVA	13,777	107
49	IANA (Internet Assigned Numbers Authority)	136	106
50	UNIVERSITA' di PADOVA	6,791	106

3.1 Activity Statistics

- Daily Visitors



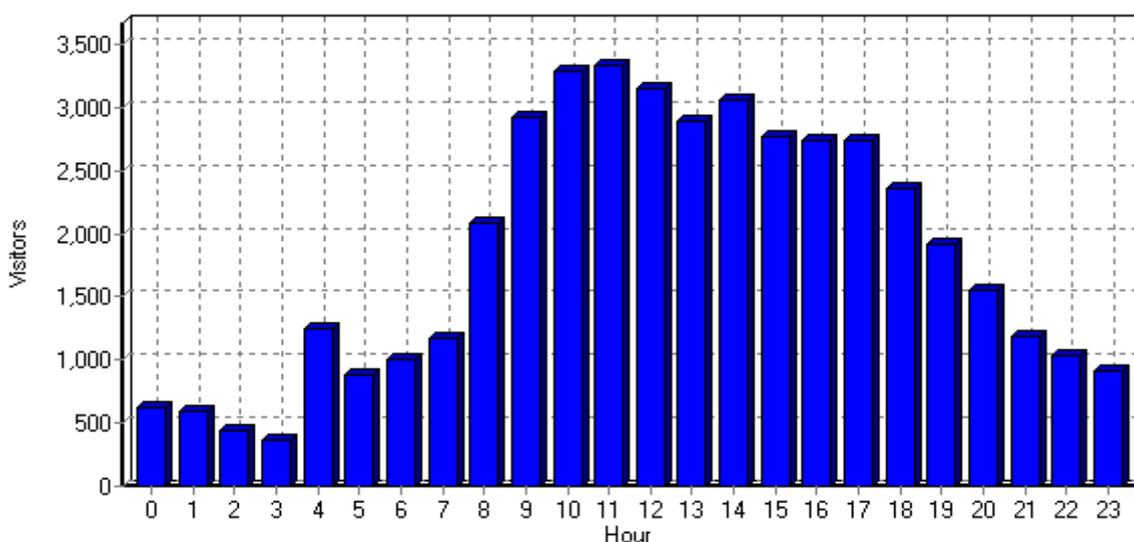
Citando dei dati, abbiamo che il totale dei visitatori è superiore ai 44000 in un mese, e che il numero medio di visitatori al giorno è di poco inferiore ai 1500.

Considerando come unità di tempo il giorno, notiamo che l'utente si collega maggiormente di lunedì, per poi andare a decrescere nel fine settimana.

Durante il periodo di esami o lauree, l'afflusso di utenza aumenta; ma quello che stupisce considerando il traffico ora per ora, è che dalle 4:00 alle 4:59 ci sono ben 1244 visitatori.

- Activity by Hour of Day

La punta massima avviene tra le 11.00 –11:59 con 3333 utenti, e la fascia che supera i 2000 è compresa tra le 8 del mattino e le 7 di sera (n.b.: i laboratori della Facoltà sono aperti dalle 8:30 alle 18:30).



Ogni visitatore in media visita circa 7 pagine e sommandole si arriva a più di 11000 al giorno, per un totale di 334789 al mese.

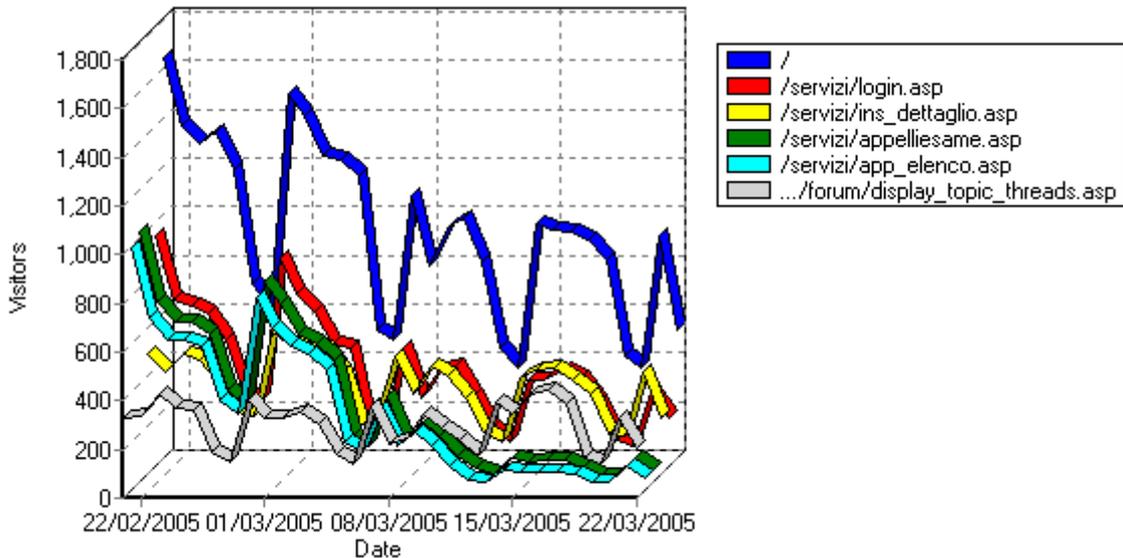
Incrociando questi dati, con quelli relativi al tempo medio speso da un utente unico su una singola pagina (Average Visit Length / Page Views) è possibile ipotizzare il comportamento-tipo dei visitatori del sito.

Nel caso specifico, una media di poche pagine viste per utente, accoppiata ad una lunga durata media delle visite registrate(10 minuti e 47 secondi), indica che i visitatori del sito trovano con relativa facilità ciò che stanno cercando e leggono a fondo i contenuti reperiti.

E' possibile predefinire la durata massima predefinita di una visita ad un sito da parte di un utente unico (Session timeout). Com'è facile comprendere, la durata di sessione è un parametro del tutto arbitrario, che nulla ha a che vedere con l'effettivo comportamento degli utenti collegati ad un sito e che può tuttavia influenzare le valutazioni del settore commerciale di un'azienda, circa la misura della fedeltà degli utenti ai siti presi in considerazione.

3.2 Access Statistics

- Pages

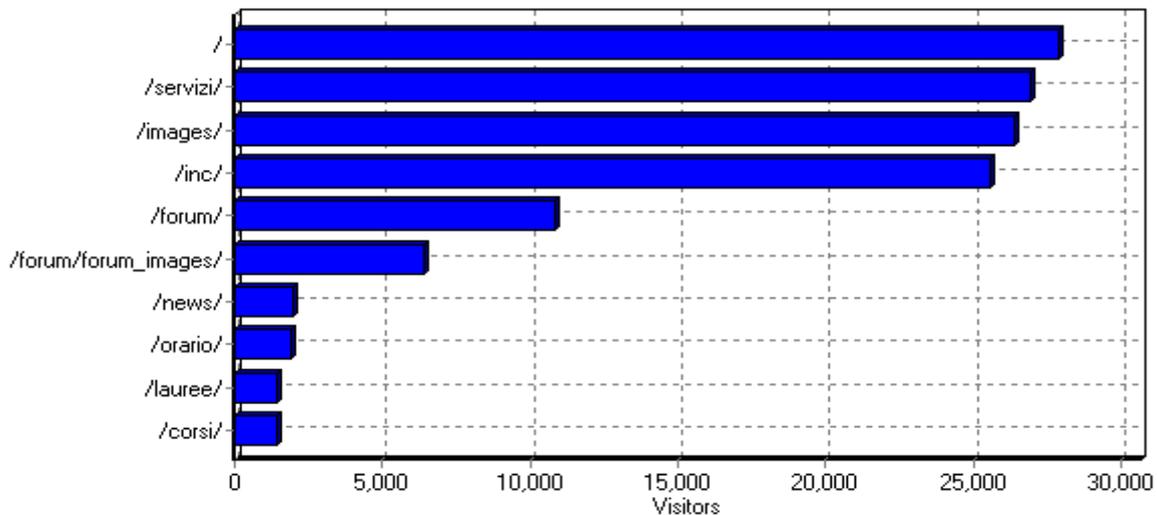


In base alle statistiche in nostro possesso le pagine più richieste, cioè quelle che hanno ricevuto più contatti in determinato arco di tempo, sono l'homepage e successivamente la directory "servizi".

Quest'informazione ci permette sia di correggere eventuali problemi di natura tecnica, come il sovraccarico del server web dovuto ad errori di programmazione (nel nostro caso, fortunatamente no); sia per risolvere problemi di struttura logica del sito: alcune pagine, potrebbero essere al vertice della classifica delle più richieste non per i loro contenuti, ma perché sono strettoie obbligate da cui passare per raggiungere determinati altri servizi.

Grazie ad un menù di link sempre presente all'interno dell'homepage, il sito della Facoltà non ha di questi problemi (nessuna richiesta incompleta).

- Directory



A proposito di directory, la classifica “Top directory”, da un’idea dell’importanza reciproca delle sezioni in cui è suddiviso un sito. La struttura logica del sito deve essere progettata in modo razionale, raggruppando i vari contenuti, in base alla loro omogeneità, sotto apposite directory.

Per quanto riguarda la Top entry pages, le pagine iniziali più richieste per ciascuna visita, compare l’homepage al vertice, salvo notare che digitando in alcuni motori di ricerca (Altavista e Lycos) o web directory (Yahoo e Supereva) le parole corsi statistica compare il sito della Biblioteca della Facoltà.

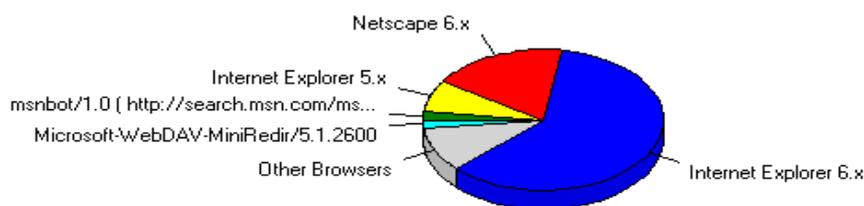
Quelle pagine che ricevono pochi contatti sono ben visibili, ma forse il loro contenuto non è sufficientemente interessante, ad esempio pagine relativo al tutorato, al campus e all’associazione Aledis.

È importante evidenziare anche la classifica delle pagine finali di una visita (Top exit pages), dove spesso la causa che spinge i visitatori a lasciare il sito dipende da una serie di collegamenti diretti ad altri siti, o

principalmente si tratta di un cattivo sviluppo dell'albero di navigazione, che finisce con il condurre gli utenti verso "pagine-imbuto" prive sia di informazioni utili sia di collegamenti verso altre sezioni del sito.

3.3 Browsers e Operating System

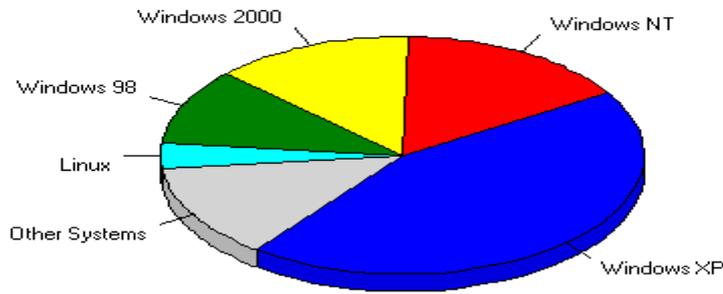
- Browsers



Per quanto concerne i browser usati, l'informazione è utile soprattutto ai responsabili tecnici, per tarare al meglio la struttura delle pagine e la presentazione dei contenuti, in modo che sono navigabili per mezzo di ciascun dei vari tipi di browser che risultano presenti nella classifica.

Analizzando l'elenco si scopre che la percentuale di visitatori utilizza Internet Explorer 6, più del 60%.

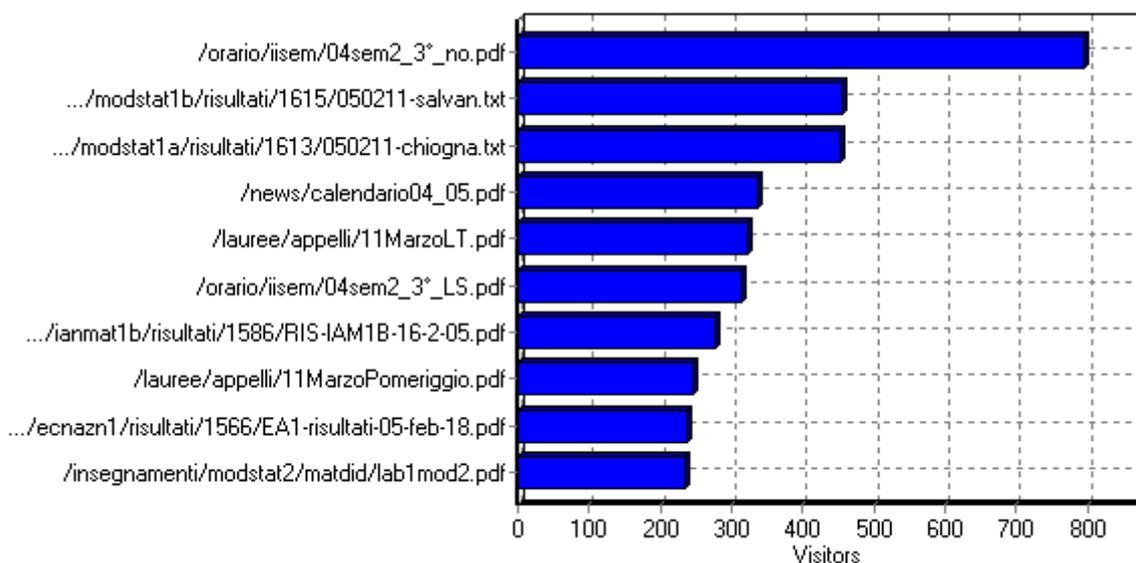
- Operating System



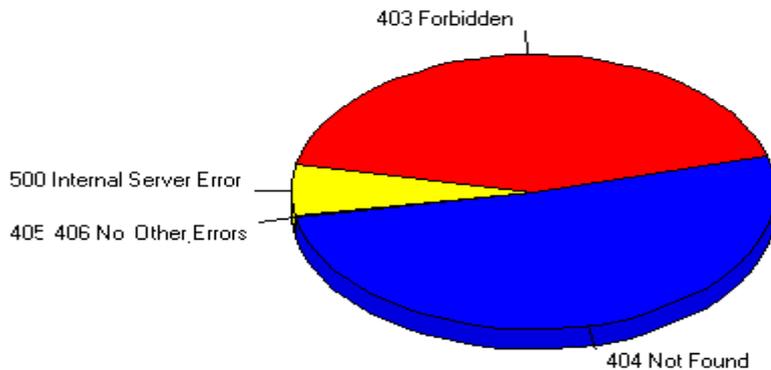
Questo ci fa capire che l'azienda Microsoft ha un monopolio anche per quanto riguarda i Sistemi operativi con windows XP in testa, con più del 44%, (Linux 3,49%). Ne consegue che quando un utente vuole scaricare un file, non deve avere problemi derivanti dalla non compatibilità tra il software e il sistema operativo che utilizza. I file scaricabili sono principalmente di estensione pdf o txt.

I visitatori in base al periodo preso in considerazione effettuano download per accedere agli orari delle lezioni, degli insegnamenti oppure ai risultati degli esami.

Downloaded Files



3.4 Errors



Considerando gli errori, la maggior parte è di tipo 'Not found', seguito dal tipo 'Forbidden', causato principalmente dalla mancata immissione dello username e password per accedere al download dei file. Solo gli utenti registrati possono visualizzare ed eventualmente scaricare i file.

3.5 Considerazioni

In definitiva, si potrebbe concludere che le rilevazioni statistiche del traffico generato da un sito, più che essere prese in se stesse come numeri assoluti, dovrebbero essere prese in senso relativo, cioè come rapporto proporzionale esistente tra le visite ricevute dalle differenti pagine del sito.

Dall'analisi di simili informazioni, i gestori di un sito possono trarre una serie di conclusioni utili a migliorare la struttura e i contenuti delle proprie pagine web, anche se rimane relativamente conoscibile il numero reale di visitatori in certo intervallo.

Di seguito viene illustrata una tabella che sintetizza gli argomenti trattati nel capitolo, ad esempio le pagine viste e i visitatori.

Hits	
Total Hits	1,887,003
Average Hits per Day	62,900
Average Hits per Visitor	42.63
Cached Requests	782,754
Failed Requests	128,084
Page Views	
Total Page Views	334,789
Average Page Views per Day	11,159
Average Page Views per Visitor	7.56
Visitors	
Total Visitors	44,260
Average Visitors per Day	1,475
Total Unique IPs	16,930
Bandwidth	
Total Bandwidth	0 B
Average Bandwidth per Day	0 B
Average Bandwidth per Hit	0 B
Average Bandwidth per Visitor	0 B

Capitolo 4: Conclusioni

Per quanto riguarda la popolarità del sito all'interno del web, a mio parere bisognerebbe effettuare una politica di cross-promotion per migliorarne la propria visibilità. La cross-promotion consiste nello scambio reciproco tra 2 siti di link testuali o di banner linkati l'uni con gli altri. L'utilità sta nel fatto che rende il sito più noto, a patto che lo scambio avvenga tra siti con lo stesso (simile) numero di visitatori. Una particolare attenzione deve essere data ai banner, i quali non devono essere mai posizionati in fondo alla pagina con sopra altri link. È importante non sottovalutare questi piccoli accorgimenti, altrimenti si rischia di rimanere al punto di partenza.

Come detto sopra, i visitatori sono importanti per capire quale sia il traffico generato dal sito, ma un altro elemento da confrontare con gli altri siti è l'argomento trattato. È consigliabile farsi linkare, ad esempio:

- da siti universitari (italiani ed europei): se desidero frequentare un corso di statistica, e sono iscritto alla Facoltà di Scienze Politiche, è utile avere il link della suddetta Facoltà sul sito di statistica e viceversa;
- da scuole superiori: orientando gli studenti ad una eventuale scelta universitaria;
- da associazioni di industriali: opportunità di lavoro, dopo gli studi. Questa unione avviene già tramite lo stage e attraverso corsi di orientamento o convegni che fanno le aziende all'interno del plesso universitario;
- da siti che forniscono servizi agli studenti: un esempio può essere dato dallo FSE (Fondo Sociale Europeo) presente nel sito. Oltre a

questo servizio si potrebbero aggiungere quello delle librerie, dove poter acquistare i libri che si possono solo consultare in biblioteca, pagandoli a prezzo scontato. Lo sconto è dovuto dal fatto che il negozio si farebbe pubblicità sul sito.

Ritornando ai visitatori, a noi interessa sapere quali tipologie di persone visita il sito. Da dati raccolti, si riscontra che sono il 10% si collega all'interno della Facoltà di Scienze Statistiche, e di conseguenza altro 90% accede al sito dal di fuori. Quest'ultimo dato risulta particolarmente interessante per due motivi:

1. non è detto che gli studenti o docenti iscritti alla suddetta Facoltà, accedano esclusivamente in sede, possono collegarsi dalla propria abitazione;
2. tolti gli individui sopra citati, è importante porsi una domanda: Come hanno fatto a trovare il sito gli altri visitatori? Generalmente utilizzano motori di ricerca o web directory, consegue che essere posizionati ai primi posti all'interno di essi sia molto utile per la propria visibilità.

Ciascun visitatore ragiona alla propria maniera, il sito riesce ad indirizzare nella giusta maniera l'utente (vedi tabella introduttiva). Da notare che ci sono delle anomalie relative ai "corsi di statistica": invece di andare sulla pagina dei corsi, va sul sito della biblioteca. Questo errore è dovuto dal fatto che il motore di ricerca considera i libri in funzione dei corsi (e materiali relativi). Al contrario se si digita "master statistica", i motori aprono la pagina giusta.

A tal proposito, i servizi che offre la Facoltà sono tutti attivi, senza problematiche. Per migliorare la chiarezza e la semplicità del sito è stato creato un menù dei servizi che rimane sempre visibile durante la navigazione. Nel sorgente HTML relativo al menù si nota che è stata utilizzata una tabella e non un frame. Quest'ultimo può creare problemi nella navigabilità.

Per quanto concerne all'aspetto esteriore, di seguito viene illustrata l'homepage.



La grafica è semplice, ma allo stesso tempo efficace. La totalità delle immagini sono statiche, forse qualche figura animata relativa ad esempio al forum o alle comunicazioni renderebbe il sito più accattivante. Premetto che sono più importanti i contenuti, ma anche l'occhio vuole la sua parte. Un punto di forza è la diversa tipologia di colori, in modo da evidenziare le varie aree.

Abbiamo accennato ai contenuti, i quali devono essere possibilmente corposi ben strutturati e impaginati all'interno del sito. I testi sulle pagine vanno scritti con attenzione e bisogna assicurarsi che siano ben attinenti ai temi trattati dal sito. Il sito della Facoltà non contiene pagine carenti di testo, e permette soprattutto all'utente di sapere in tempo reale il percorso che ha effettuato. Esempio: home > Santa Caterina: il campus

Sempre riguardo ai contenuti, per un buon posizionamento è utile puntare sulla qualità, non sulla quantità. Trovarsi di fronte un sito che sperimenta e pensa a nuove soluzioni per migliorare il proprio sito, avrà sempre una maggior visibilità all'interno della rete. Da notare che in precedenza il sito consentiva a chiunque di accedere ai servizi senza effettuare la registrazione: il risultato era che la maggior parte degli utenti iscritti agli esami aveva nomi di fantasia. Più recente è il vincolo nel forum, dove alcune persone lo usavano solo per insultare.

La tendenza per il futuro sta già tutta nell'analisi della situazione attuale con due o tre motori di ricerca attorno ai quali ruota la maggior parte del traffico web. È importante che il sito della Facoltà rimanga nelle prime tre pagine di questi tre siti per ottenere una notevole visibilità. La visibilità si crea anche con il "passaparola", in altre parole un utente soddisfatto torna e lo consiglia ad altri.

Al giorno d'oggi i trucchi non servono, anzi fanno una pessima pubblicità al sito; perciò come espresso in precedenza bisogna puntare sulla qualità dei servizi che l'utente richiede.

Può sembrare un controsenso, ma il sito deve essere fatto in funzione dell'utente, e non del motore di ricerca, infatti, la Facoltà di Scienze Statistiche si mette al servizio degli utenti, siano essi studenti o docenti.

Capitolo 5: Bibliografia

5.1 Libri e articoli

- Libri:

Vaccaio Silvia, *Ricerca su Internet*, Mondadori Informatica, 2005

Marangoni Roberto -Cucca Alessandro, *Motori di ricerca*, Hoepli, 2001

Lupi Mauro, *Motori di ricerca e visibilità sul Web*, Apogeo, 2001

- Articoli di Chiantia Danilo:

Dove posizionare le parole chiave,

Differenza tra indicizzazione e posizionamento.

In questi articoli sono citati due libri:

Sevenoaks R., *Posizionamento sui motori di ricerca*,

Mirando L., *La terza generazione dei motori di ricerca.*

- Siti internet:

<http://ad.bloogz.com> (link popularity, page rank)

<http://www.newbie.it> (spiders, motori di ricerca, posizionamento)

<http://www.mitama.it/marketing> (cross promotion, scambio banner)

<http://www.motoriricerca.info> (glossario)

Tra parentesi sono elencati gli argomenti ricavati dai siti: il periodo di visualizzazione e scaricamento sono stati effettuati nel maggio del 2005.

5.2 Dati e relative tabelle

I dati sono stati gentilmente concessi dalla Facoltà di Scienze Statistiche, nella persona di Mirko Moro.

Le principali tabelle utilizzate per l'indagine statistica:

- Daily Activity

Date	Hits	Page Views	Visitors	Average Visit Length	Bandwidth (KB)
Mon 2/21/2005	120,316	20,390	2,264	10:24	0
Tue 2/22/2005	98,936	16,118	1,959	10:24	0
Wed 2/23/2005	99,925	18,190	1,991	12:59	0
Thu 2/24/2005	94,387	16,570	1,991	10:47	0
Fri 2/25/2005	80,619	13,140	1,723	09:57	0
Sat 2/26/2005	43,271	8,251	1,260	09:14	0
Sun 2/27/2005	37,352	6,979	1,047	09:22	0
Mon 2/28/2005	122,618	22,437	2,004	12:43	0
Tue 3/1/2005	112,650	22,841	1,960	15:02	0
Wed 3/2/2005	95,535	16,976	1,852	11:43	0
Thu 3/3/2005	85,966	14,575	1,896	10:09	0
Fri 3/4/2005	78,631	12,666	1,686	09:59	0
Sat 3/5/2005	28,585	5,831	1,007	13:45	0
Sun 3/6/2005	27,589	5,179	1,089	08:58	0
Mon 3/7/2005	76,242	13,160	1,707	09:36	0
Tue 3/8/2005	55,087	10,860	1,667	09:46	0
Wed 3/9/2005	63,212	10,939	1,589	10:37	0
Thu 3/10/2005	64,503	11,356	1,513	10:38	0
Fri 3/11/2005	52,025	8,931	1,343	10:47	0
Sat 3/12/2005	26,287	5,351	870	10:36	0
Sun 3/13/2005	18,687	3,664	714	10:13	0
Mon 3/14/2005	63,674	9,976	1,400	09:52	0
Tue 3/15/2005	55,979	9,987	1,424	11:43	0
Wed 3/16/2005	58,436	11,002	1,492	11:16	0
Thu 3/17/2005	57,655	10,342	1,338	11:13	0
Fri 3/18/2005	42,269	7,653	1,325	10:10	0
Sat 3/19/2005	20,431	3,854	879	09:13	0
Sun 3/20/2005	17,157	3,016	787	08:44	0
Mon 3/21/2005	54,860	8,832	1,557	08:49	0
Tue 3/22/2005	34,119	5,723	926	10:32	0
Total	1,887,003	334,789	44,260	10:47	0

- Activity by Hour of Day

Hour	Hits	Visitors	Bandwidth (KB)
00:00 - 00:59	15,000	629	0
01:00 - 01:59	8,432	597	0
02:00 - 02:59	5,189	436	0
03:00 - 03:59	2,703	371	0
04:00 - 04:59	7,258	1,244	0
05:00 - 05:59	3,615	886	0
06:00 - 06:59	3,772	1,005	0
07:00 - 07:59	14,314	1,176	0
08:00 - 08:59	72,959	2,089	0
09:00 - 09:59	131,203	2,915	0
10:00 - 10:59	184,596	3,289	0
11:00 - 11:59	172,840	3,333	0
12:00 - 12:59	174,951	3,153	0
13:00 - 13:59	156,776	2,887	0
14:00 - 14:59	148,917	3,061	0
15:00 - 15:59	144,329	2,771	0
16:00 - 16:59	149,768	2,731	0
17:00 - 17:59	130,948	2,737	0
18:00 - 18:59	113,652	2,361	0
19:00 - 19:59	76,066	1,912	0
20:00 - 20:59	58,711	1,547	0
21:00 - 21:59	45,204	1,180	0
22:00 - 22:59	37,938	1,040	0
23:00 - 23:59	27,862	910	0

- Browsers

	Browser	Hits	Visitors	% of Total Visitors
1	Internet Explorer 6.x	1,337,866	23,478	60.26%
2	Netscape 6.x	293,112	7,130	18.30%
3	Internet Explorer 5.x	169,890	2,852	7.32%
4	msnbot/1.0 (http://search.msn.com/msnbot.htm)	4,699	774	1.99%
5	Microsoft-WebDAV-MiniRedir/5.1.2600	8,276	629	1.61%
6	Microsoft Data Access Internet Publishing Provider Cache M.	978	528	1.36%
7	contype	917	400	1.03%
8	Microsoft Data Access Internet Publishing Provider DAV 1.1	768	386	0.99%
9	Netscape 7.x	11,397	362	0.93%
10	Netscape 4.x	750	328	0.84%

- Operating System

	Operating System	Hits	Visitors	% of Total Visitors
1	Windows XP	771,123	15,408	44.17%
2	Windows NT	412,269	5,495	15.75%
3	Windows 2000	360,223	4,642	13.31%
4	Windows 98	181,641	3,628	10.40%
5	Others	33,001	3,068	8.79%
6	Linux	46,601	1,216	3.49%
7	Windows ME	41,982	983	2.82%
8	Mac OS	7,426	305	0.87%
9	Windows 95	3,025	119	0.34%
10	Windows Server 2003	689	18	0.05%
11	Sun OS	95	4	0.01%
12	OSF1	8	1	0.00%

- Errors

	Code	Description	Count
1	404	Not Found	65,394
2	403	Forbidden	54,698
3	500	Internal Server Error	7,838
4	405	Method Not Allowed	98
5	406	Not Acceptable	24
6	400	Bad Request	15
7	416	Requested Range Not Satisfiable	14
8	501	Not Implemented	3

5.3 Glossario

Anchor text (testo dell'ancora)

Il testo che compone un link testuale, ovvero quello compreso tra il tag HTML <a> (ancora) di apertura e il rispettivo tag di chiusura (). Le parole contenute in un link testuale possono contribuire a tematizzare la pagina verso la quale il link punta.

Backlink

Dal punto di vista di un sito web, un backlink è un link ricevuto da un altro sito. Dal punto di vista di una singola pagina web, un backlink è un link ricevuto da altre pagine, dello stesso sito o di siti esterni.

Banned (bandito)

Un sito web è "bandito", ovvero cancellato, dagli archivi di un motore di ricerca quando il motore individua nel sito l'uso di tecniche scorrette (spam) volte ad aumentare artificialmente la sua posizione sui risultati delle ricerche. Si tratta del tipo di penalizzazione più grave che un motore può infliggere.

Click-through / Click-thru

Un click-thru corrisponde ad un click effettuato da un utente sul link ad una pagina web riportata dal motore in seguito ad una ricerca. Alcuni motori effettuano statistiche sulle pagine più cliccate dagli utenti e usano questa informazione per far salire di posizione le pagine maggiormente cliccate.

Cloaking

È una tecnica che permette di variare "al volo" i contenuti di una pagina web, a seconda di chi la visita. Ad esempio, è possibile mostrare certi testi agli utenti e altri testi allo spider di un motore di

ricerca. I motori di ricerca solitamente non amano i siti che usano il cloaking per mostrare agli spider contenuti diversi da quelli mostrati agli utenti; tale tecnica è infatti considerata spam e i siti che ne fanno uso possono venire banditi dai motori.

Crawler

Programma usato dai motori di ricerca per scandagliare il web e individuare le pagine da prelevare. È quasi sinonimo di "spider", in quanto il suo fine è leggermente diverso.

Contenuti dinamici

Sono i contenuti di una pagina web che variano al cambiare di parametri vari. I motori di ricerca non hanno problemi ad analizzare questi contenuti, se vengono prodotti da un linguaggio "lato server" (PHP, ASP, ecc.) ma incontrano problemi quando la pagina fa uso di linguaggi "lato cliente" (Javascript in testa), in quanto gli spider che prelevano le pagine non sono in grado di interpretare questi linguaggi.

Deep crawl (scansione profonda)

È un grande processo di scansione del web da parte degli spider dei motori di ricerca: consiste nel visitare la maggior parte (o tutti) i siti presenti nel proprio archivio e a prelevare il massimo numero di pagine possibile.

Densità parole-chiave (keyword density)

La densità indica quanto frequentemente una parola-chiave appare in un testo. Questo indice di frequenza è solitamente indicato come percentuale della parola-chiave sull'intero testo. In un'interpretazione leggermente diversa, la percentuale può essere riferita alle sole parole del testo che si ripetono almeno una volta.

Secondo questa seconda interpretazione, dunque, una parola che appare una sola volta nel testo avrebbe una densità pari a zero. Quando si progetta una pagina a tema, su un argomento specifico, e si individuano le principali parole-chiave sulle quali focalizzare il testo, è bene includerle in percentuale non eccessiva, possibilmente non superiore al 10%. Ovviamente non si tratta di una regola ferrea: l'intento rimane comunque quello di mantenere il testo leggibile, privo di eccessive ripetizioni dei termini.

Descrizione

Testo di una pagina web che ne descrive i contenuti e gli argomenti trattati. Alcuni motori di ricerca estrapolano la descrizione di una pagina dai contenuti del meta tag DESCRIPTION, nato proprio per questo scopo.

Altri motori creano "al volo" una descrizione estrapolando il testo dalle zone della pagina che contengono le keyword cercate dall'utente. Se una pagina è presente in ODP, alcuni motori possono mostrare la descrizione che gli editor della directory hanno assegnato alla pagina.

Directory

Catalogo di siti web ordinato per categoria. Ogni sito web viene ospitato in una categoria relativa ai contenuti del sito. Generalmente le directory, a differenza dei motori di ricerca "propriamente detti" selezionano i siti web proposti dagli utenti manualmente, includendoli nel catalogo solo se vengono ritenuti interessanti e di qualità.

Doorway page (pagina doorway; dette anche: Gateway page)

Pagina i cui contenuti testuali sono strettamente focalizzati su una parola-chiave o su una breve frase-chiave. Nell'uso comune, il termine ha assunto una connotazione negativa, in quanto generalmente le doorway non vengono usate per ospitare contenuti testuali appetibili ai visitatori, ma solo parole chiave e link studiati per aumentare artificialmente la rilevanza della pagina agli occhi dei motori di ricerca.

Everflux

Fenomeno che vede un sito cambiare repentinamente e ripetutamente posizione nei risultati delle ricerche di Google. A distanza di anche un solo giorno, una pagina può apparire in posizioni molto diverse o addirittura scomparire temporaneamente dall'archivio del motore di ricerca. Tale effetto è causato dal fatto che Google può attribuire particolare importanza alle pagine nuove o a quelle che sono state modificate di recente, premiandole con una posizione leggermente migliore. Tali pagine nuove o con nuovi contenuti, dunque, possono apparire in una posizione prominente per un giorno, per poi scendere di posizione nei giorni successivi, quando i contenuti non vengono più considerati "freschi".

FFA (Free For All) link

Pagine web in cui qualunque visitatore può aggiungere un link al proprio sito. Sono nate per incrementare artificialmente la Link Popularity di un sito e solitamente consistono in enormi pagine piene esclusivamente di link e prive di alcun contenuto appetibile ai visitatori. Sono estremamente mal viste dai motori di ricerca che

spesso adottano soluzioni per ignorarle e in alcuni casi per penalizzare i siti web che ne fanno uso.

Frameset / Frame / Iframe

Il Frameset è una tecnica che permette di dividere lo spazio della finestra di visualizzazione del browser in zone rettangolari, chiamate "frame", ognuna delle quali mostra un file HTML differente. Il metodo viene usato per far caricare una sola volta agli utenti quelle zone della finestra destinate a rimanere fisse e a non cambiare di contenuto (una intestazione in alto, un menu di navigazione a lato, ecc.). È consigliato usare i frame solo quando strettamente indispensabili, in quanto possono nascere problemi di usabilità e di indicizzazione del sito con i motori di ricerca.

Google Dance

Periodo di tempo che Google dedica ad aggiornare il proprio database principale. Durante l'aggiornamento Google ricalcola i valori di PageRankTM di tutte le pagine in archivio. La Google Dance avviene solitamente una volta al mese e dura alcuni giorni; viene chiamata "danza" in quanto durante tale periodo le posizioni dei siti su Google possono cambiare ripetutamente, apparendo spesso inesatte e fuorvianti. Tutto si assesta quando l'aggiornamento è terminato.

In più di un sito web italiano, abbiamo letto affermazioni secondo le quali durante la Google Dance sarebbe possibile sapere in anticipo i risultati dell'aggiornamento di Google, ma ciò è del tutto errato; al contrario, durante la "dance" non bisognerebbe fare affidamento ai risultati di qualunque datacenter di Google, che appaiono instabili e incorretti.

Keyword (parola-chiave)

Termine particolarmente attinente ad un dato argomento. Gli utenti che effettuano ricerche sui motori inseriscono spesso una o più parole-chiave che ritengono inerenti ai contenuti che intendono trovare. I motori di ricerca riportano poi all'utente un elenco di pagine web in cui compaiono le parole-chiave inserite.

Keyphrase (frase-chiave)

Insieme di poche parole-chiave (solitamente due o tre) che formano una breve frase o un'espressione particolarmente attinente ad un dato argomento.

Inbound link

Dal punto di vista di un sito, è un link ricevuto da altri siti. Gli inbound link sono estremamente utili in quanto aumentano la Link Popularity di un sito e contribuiscono ad incrementare il valore di PageRankTM calcolato da Google.

Indicizzazione

Atto compiuto dal motore di ricerca e che consiste nell'aggiungere una pagina web all'archivio dei siti (database) del motore. È sinonimo di "registrazione".

Link farm (letteralmente: fattoria di link; più usato: fabbrica di link)

Sistema o servizio attraverso il quale è possibile incrementare il numero di link ricevuti da un sito. Essendo generalmente una tecnica volta esclusivamente ad aumentare la Link Popularity di un sito, è mal vista da tutti i motori di ricerca.

Link Popularity (popolarità da link)

Il numero di link in giro per il web che puntano ad una determinata pagina o sito. La Link Popularity è uno dei fattori che i motori di

ricerca prendono in considerazione per attribuire la posizione di una pagina nei risultati delle ricerche. I motori desumono che una pagina sia più rilevante se la sua Link Popularity è alta.

Link reciproco

Avviene quando il sito A ospita nelle sue pagine un link verso il sito B ed il sito B contraccambia con un link verso il sito A. I link reciproci vengono massicciamente utilizzati dai webmaster per aumentare la Link Popularity dei propri siti web. Logica vorrebbe che col tempo i motori di ricerca daranno preferenza ai link spontanei più che ai link nati in seguito ad accordi di scambio.

Meta tag

Un tag HTML che ospita informazioni di vario genere. Alcuni di questi tag (come il DESCRIPTION e il KEYWORDS) ospitano contenuti dedicati ai motori di ricerca.

Motore di ricerca

Strumento per mezzo del quale è possibile reperire informazioni specifiche all'interno di una grande mole di documenti. L'utente effettua una ricerca inserendo alcuni termini che, presumibilmente, vengono usati in testi trattanti l'argomento ricercato ed il motore restituisce un elenco di pagine web e altri tipi di documenti che contengono il termini inseriti.

ODP: Open Directory Project

Directory nata e cresciuta attraverso la cooperazione di migliaia di volontari, che recensiscono continuamente siti web con la finalità di costruire una directory con tanti contenuti di qualità. Il database dei siti ospitati dalla directory è a disposizione di chiunque e viene usato, tra gli altri, anche da Google.

Ottimizzazione

Processo mediante il quale i contenuti e le caratteristiche tecniche di una pagina o di un intero sito web vengono corretti e migliorati con la finalità di fornire contenuti più appetibili ai visitatori e conseguire posizioni migliori sui motori di ricerca.

Outbound link

Link verso l'esterno, ovvero che punta alla pagina di un altro sito web.

PageRank™

Valore che Google attribuisce ad ogni pagina web contenuta nel proprio archivio, più il valore di PageRank™ di una pagina è alto e più la pagina sale di posizione nelle ricerche. Tale valore è attribuito basandosi sulla quantità e la "qualità" dei link ricevuti dalla pagina; la qualità di un link aumenta con l'incrementare del valore di PageRank™ della pagina in cui esso risiede.

Pagina orfana (orphan page)

Pagina web che non riceve link da nessun'altra pagina. Generalmente questo tipo di pagine non riesce ad entrare negli archivi dei motori di ricerca, in quanto non rientra nel modello richiesto dai motori: ogni pagina deve ricevere almeno un link da un'altra pagina.

PFI: Pay For Inclusion (pagare per l'inclusione)

Servizio commerciale offerto da alcuni motori che garantisce l'inserimento di una o più pagine web nel proprio archivio dei siti solo a fronte del pagamento di una cifra in denaro. I motori più grandi, come Google o Fast, non offrono questo tipo di servizio in quanto preferiscono includere nei propri archivi il massimo numero

di pagine possibile, indipendentemente dalla disponibilità finanziaria dei proprietari dei siti web.

Posizionamento

Insieme di tecniche e strategie volte a migliorare la posizione di un sito web nei risultati delle ricerche sui motori. Le tecniche possono essere più o meno etiche secondo il "temperamento" del SEO.

Prominenza

Letteralmente "vicinanza", viene usato per riferirsi alla vicinanza reciproca tra due o più parole-chiave, oppure alla vicinanza di una parola-chiave ad un punto particolare di una pagina web (una intestazione, l'inizio del primo paragrafo di testo, ecc.). La prominenza delle keyword ai "punti chiave" di una pagina è uno dei fattori che molti motori di ricerca prendono in considerazione per assegnare un valore di rilevanza tra la pagina e i termini cercati dall'utente.

Redirect (redirezione)

Accorgimento tecnico usato su alcune pagine web che dirotta automaticamente il visitatore su un'altra pagina. I motori di ricerca non amano le pagine che usano i redirect, in quanto vengono spesso utilizzate per nascondere soluzioni "poco pulite", quali le pagine doorway.

Registrazione

Atto compiuto dal motore di ricerca e che consiste nell'aggiungere una pagina web all'archivio dei siti (database) del motore. È sinonimo di "indicizzazione".

Risultati organici (organic results)

Quella parte dei risultati di una ricerca che è frutto esclusivamente degli algoritmi di rilevanza applicati dal motore e che non è influenzata dalle inserzioni a pagamento o pubblicitarie.

Robot

Software che effettua in automatico l'operazione di visita di un sito e prelievo delle sue pagine web. È un termine simile a "spider", ma più generico, in quanto non strettamente relativo ai motori di ricerca.

Robots.txt (file)

File di testo, da inserire nella directory principale di un dominio, che contiene al suo interno alcune direttive che molti spider dei motori di ricerca si sono impegnati a seguire. Viene usato per impedire ai motori di ricerca di scandagliare e prelevare file o directory che non si vuole siano archiviati dai motori.

Segnalazione

Operazione volta a fornire al motore di ricerca l'indirizzo (URL) di una pagina web da visitare. La maggior parte dei motori parte dalla pagina segnalata per individuare e archiviare automaticamente tutte le pagine di un sito web. Altri motori (come Altavista) richiedono che le pagine di un sito vengano segnalate una per una.

SEO: Search Engine Optimizer (ottimizzatore per motori di ricerca)

La figura professionale che si occupa di ottimizzare i siti e posizionarli sui motori.

SERP: Search Engine Results Page (pagine di risultati della ricerca)

La pagina di risultati che i motori restituiscono e che contiene l'elenco delle pagine web inerenti la ricerca effettuata.

SPAM / Web SPAM / Spamming / Spamdexing

Qualunque tecnica volta a fuorviare i motori di ricerca circa i contenuti di un sito, la sua attinenza ad alcuni temi specifici o la sua popolarità sul web. SPAMMER è colui che adotta tali tecniche, in italiano viene anche usato il (brutto) neologismo "SPAMMARE", che indica l'azione dello spammer. SPAM è anche il termine che generalmente viene utilizzato per indicare la posta indesiderata che affolla le caselle postali, pertanto bisogna fare attenzione a non far confusione con i due diversi significati.

Spider

Software che effettua in automatico l'operazione di visita di un sito e prelievo delle sue pagine web. È un termine simile al più generico "robot", ma maggiormente usato quando il software è usato espressamente da un motore di ricerca.

Stemming

Linguisticamente parlando, è l'atto di estrarre da una parola la sua radice. Alcuni (pochi) motori di ricerca non si limitano a cercare le pagine che contengono esattamente le keyword immesse dall'utente ma si spingono oltre fino a prendere in considerazione anche quei termini che dividono con quelli immessi la medesima radice. Ad esempio la ricerca della parola "mangiato" potrebbe essere ampliata anche alle parole "mangiati", "mangiate", "mangiare", "mangiando", ecc.

Testo nascosto

Testo di una pagina che, attraverso alcuni accorgimenti tecnici, non appare agli utenti ma risulta invece ben visibile al motore di ricerca. È uno dei trucchi più diffusi (e più penalizzati) per inserire molte parole chiave in una pagina senza pregiudicare la sua estetica agli occhi dei visitatori.

Welcome page (pagina di benvenuto)

Come le pagine doorway è una pagina web i cui contenuti testuali sono focalizzati su un argomento ben preciso e che fa un uso ampio, ma non eccessivo, delle parole chiave collegate al tema trattato. A differenza delle pagine doorway, la pagine di benvenuto sono dedicate anche agli utenti e presentano contenuti testuali realmente appetibili ai visitatori. I motori di ricerca apprezzano particolarmente questo genere di pagine tematizzate.