

Introduzione ai motori di ricerca

Cosa sono i motori di ricerca

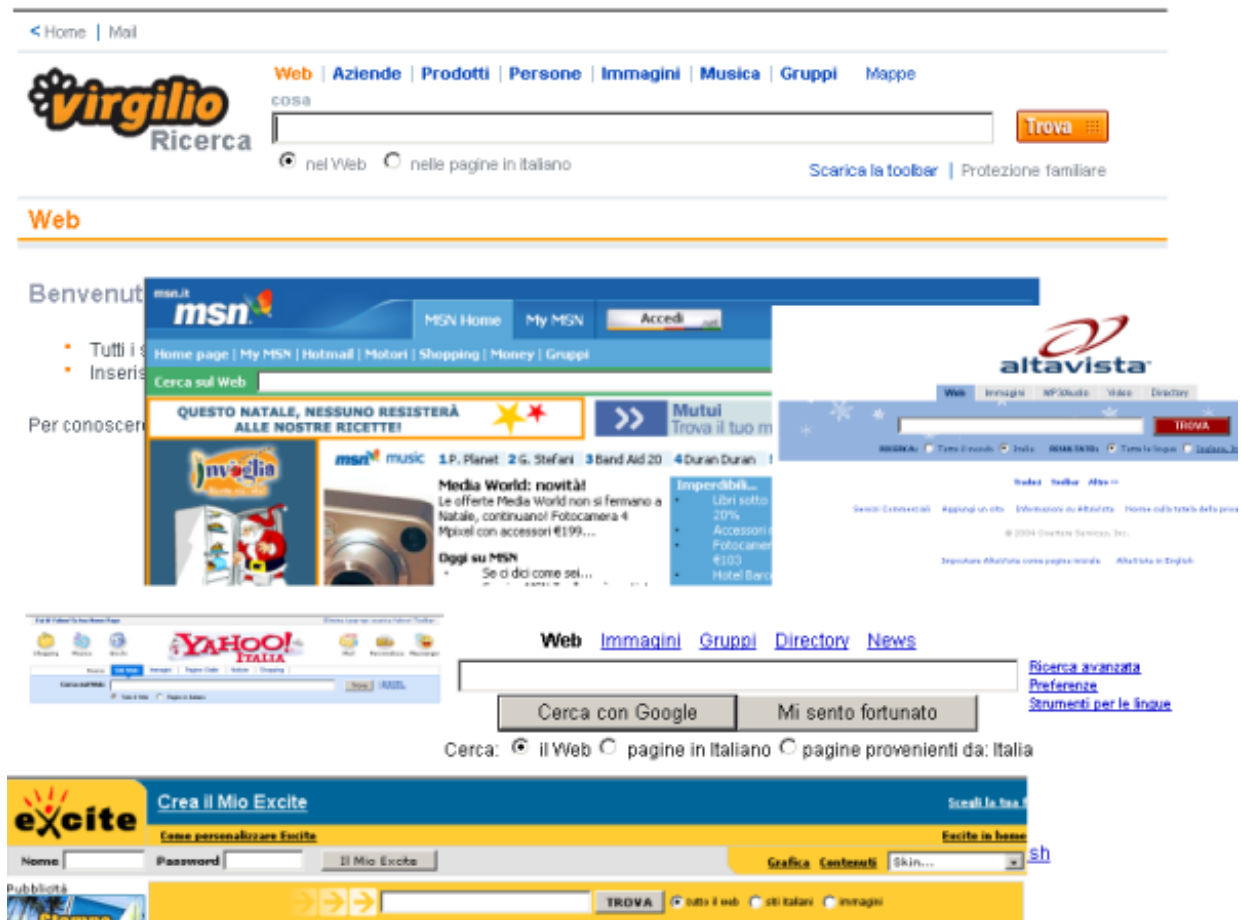
L'enorme quantità di siti web e di documenti presenti in Internet fa sì che trovare ciò che si sta cercando non sia semplicissimo.

Se da una parte può essere utile avere una serie di siti preferiti di riferimento in cui rintracciare le informazioni desiderate, dall'altra esplorare a caso la rete per trovare nuovi fonti di informazione può essere un'operazione estremamente lenta e frustrante.

I motori di ricerca permettono di risolvere questa situazione. Si tratta infatti di strumenti che consultano l'intera rete mondiale memorizzando tutti i contenuti e indicizzandoli. In pratica, se immaginiamo il web come un'enorme biblioteca (in cui peraltro chiunque o quasi può aggiungere nuovi documenti), i motori di ricerca svolgono il ruolo di sportello informazioni dicendo ai richiedenti dove trovare il volume desiderato. In considerazione delle evidenti differenze rispetto alle biblioteche, il meccanismo è leggermente più complesso, ma l'analogia può rendere l'idea.

Attualmente un motore come Google, indicizza circa più di 4 miliardi di pagine, è cioè molto vicino a conoscere il contenuto di tutte le pagine pubbliche presenti. Altri motori hanno valori più bassi, nell'ordine di centinaia di milioni di pagine. La differenza è data principalmente dal fatto che molti motori non sono interessati a conoscere tutte le pagine di un sito, ma si limitano alla home page e poche altre.

Usare un motore di ricerca è molto semplice. Basta inserire nell'apposito spazio una o più parole che indichino l'argomento cercato. In questo modo si otterrà come risposta una serie di siti che rispondono al criterio specificato e quindi consultare direttamente quelli che sembrano essere più interessanti e utili.



Motori e Directory

Con il termine motore di ricerca si tende generalmente ad indicare i siti che permettono di effettuare ricerche sul web, ma a ben guardare questi strumenti possono essere distinti tra motori di ricerca veri e propri e directory (indici)

I motori di ricerca in senso stretto sono quelli che rispondono più precisamente alla descrizione di cui sopra: in base ai termini specificati restituiscono un elenco di siti che contengono i termini cercati. Il meccanismo di identificazione delle parole usate per la ricerca all'interno delle pagine memorizzate dal motore varia da motore a motore, e anche l'elenco risultante può essere ordinato in modi differenti a seconda dei criteri utilizzati dai progettisti dei motori. Il sistema di produzione di risultati ordinati viene definito Ranking.

Le directory hanno invece un proprio criterio di catalogazione al cui interno vengono inseriti i siti. Si tratta quindi di un repertorio basato sui contenuti dei siti che permette di effettuare ricerche non tanto sulla base di termini chiave, quanto sulla categoria. In questo modo, dopo avere selezionate l'argomento principale e uno o due sottocategorie, si giunge ad un elenco in cui sono raggruppati tutti i siti affini.

In effetti alcune directory permettono anche di eseguire ricerche nella modalità tipica dei motori di ricerca. Questo alle volte genera qualche confusione nel percepire la differenza tra le due tipologie.

Le directory sono comunque facilmente riconoscibili perchè offrono bene in vista la categorie dei propri contenuti; ad esempio Arte e cultura, che si differenzia poi in Architettura e design, Arti dello spettacolo, Arti visive, Musica e altre, o Sport e tempo libero, che contiene anche le sottocategorie Auto e moto, Casa e giardino, Cucina, Sport, Viaggi.

d m o z open directory project

[A proposito di dmoz](#) | [denuncia abuso/spam](#) | [aiuto](#)

Cerca in tutta la directory

[Top: World: Italiano \(145.313\)](#) [Descrizione](#)
[FAQ](#)

- [Junior@ \(190\)](#)
- [Regionale \(63.861\)](#)

- [Acquisti \(2.269\)](#)
- [Affari \(22.487\)](#)
- [Arte \(17.225\)](#)
- [Casa \(521\)](#)
- [Computer \(5.714\)](#)
- [Consultazione \(4.238\)](#)
- [Giochi \(2.003\)](#)
- [Notizie \(357\)](#)
- [Salute \(3.385\)](#)
- [Scienza \(4.622\)](#)
- [Società \(8.344\)](#)
- [Sport \(5.483\)](#)
- [Tempo Libero \(4.804\)](#)

- [Italia@ \(60.637\)](#)
- [San Marino@ \(144\)](#)
- [Svizzera@ \(1.032\)](#)
- [Unione Europea@ \(84\)](#)
- [Vaticano@ \(18\)](#)

Una importante differenza tra motori e directory è legata al modo automatico o meno di registrazione dei contenuti. Nel caso dei motori di ricerca appositi agenti automatici percorrono incessantemente il web per trovare contenuti che vengono poi memorizzati nei

server e subito resi disponibili senza alcuna ulteriore valutazione di merito. Nel caso delle directory l'inserimento in una particolare categoria è generalmente decisa da un operatore umano che quindi seleziona, decide il posizionamento e descrive con parole proprie il sito inserito.

Usare i motori di ricerca

Esamineremo ora i principali usi dei motori di ricerca, o perlomeno quelli di uso più comune o legati all'attività del webmaster.

Consultazione

Nel caso della semplice consultazione, ovvero della ricerca di informazioni sulla base di parole chiave, è opportuno sottolineare che le ricerche possono essere effettuate usando uno o più termini, che possono essere usati in modalità AND (devono essere tutti presenti nella pagina cercata) o in modalità OR (è sufficiente che sia presente almeno uno dei termini cercati). Spesso una ricerca più precisa, o ristretta, è utile al fine di evidenziare un numero limitato di pagine ed evitare di ottenere un elenco di centinaia di siti in risposta.

Analogamente è generalmente possibile specificare intere frasi che devono essere rintracciate integralmente. In questo modo la ricerca è ancora più specifica.

E' inoltre utile effettuare la ricerca direttamente sul sito nazionale del motore di ricerca, o specificando in quale lingua si desiderano i risultati. Anche questo accorgimento permette di limitare il numero di risposte e di ottenere risultati più specifici.

In linea generale è preferibile fare una ricerca partendo da termini specifici. Nel caso non si trovi ciò che si stava cercando si può allargare la ricerca usando termini più generici.

Ogni motore ha le proprie regole sintattiche e di funzionamento. E' quindi opportuno consultare attentamente quali meccanismi sono utilizzati dal motore in uso. Ad esempio alcuni motori considerano per default i termini in OR, mentre altri in AND.

Inoltre i motori di ricerca dispongono spesso di una funzione di ricerca avanzata tramite la quale specificare criteri particolari, quali l'esclusione di termini dalla ricerca, il numero di risultati elencati per pagina, quali domini includere ecc.

Ricerca avanzata [Suggerimenti per la ricerca](#) | [Tutto su Google](#)

Trova risultati	che contengano tutte le seguenti parole	<input type="text"/>	10 risultati
	che contengano la seguente frase	<input type="text"/>	<input type="button" value="Cerca con Google"/>
	che contengano una qualunque delle seguenti parole	<input type="text"/>	
	che non contengano le seguenti parole	<input type="text"/>	
Lingua	Visualizza solo le pagine scritte in	<input type="text" value="tutte le lingue"/>	
Formato file	<input type="button" value="Visualizza"/> i risultati in	<input type="text" value="qualsiasi formato"/>	
Data	Visualizza le pagine Web aggiornate riferite a	<input type="text" value="tutto l'archivio"/>	
Cerca in	Visualizza i risultati quando i termini selezionati appaiono	<input type="text" value="in una qualsiasi parte nella pagina"/>	
Domini	<input type="button" value="Visualizza"/> i risultati contenuti nel seguente sito o dominio	<input type="text" value="Esempi: .org, google.com"/> Ulteriori informazioni	

Cerca pagina specifica

Pagine simili	Cerca pagine simili alla pagina	<input type="text"/>	<input type="button" value="Cerca"/>
		Esempio: www.google.com/help.html	
Collegamenti	Cerca pagine che contengano collegamenti alla seguente pagina	<input type="text"/>	<input type="button" value="Cerca"/>

Registrazione

Al fine di avere il proprio sito disponibile nei motori di ricerca è spesso necessario registrarsi, ovvero indicare al motore stesso l'esistenza del sito.

In generale i motori di ricerca prevedono un'apposito modulo di segnalazione tramite il quale è possibile comunicare la propria esistenza ed essere messi in lista per essere visitati. I tempi di acquisizione sono variabili e possono essere di pochi giorni o di svariate settimane. I tempi più lunghi sono spesso previsti dalle Directory per la loro componente di verifica umana.

Occorre però specificare che molti motori di ricerca eseguono comunque l'indicizzazione di un nuovo sito se è in qualche modo linkato da altri siti, la registrazione può essere una sicurezza o rendere più veloce il processo.

In altri casi la presenza in un motore di ricerca è subordinata ad un'iscrizione a pagamento. E' questa una tendenza presente da alcuni mesi a questa parte che sembra andare in controtendenza rispetto alla logica dei motori stessi, che per essere affidabili e di valore devono essere in grado di segnalare qualunque sito risponda alla ricerca impostata.

Questo fenomeno mostra però che tipo di importanza abbiano assunto di recente i motori di ricerca. Non essere censiti significa non esistere nel web, di conseguenza anche la semplice presenza in un motore acquista valore e diventa un servizio che può essere venduto.

Google Aggiunta/Rimozione di un URL

[Home](#)
[Tutto su Google](#)
[Pubblicità su Google](#)
[Informazioni per webmaster](#)
Aggiunta/Rimozione di un URL
[Aggiungere Google al vostro sito](#)

Aggiunta/Aggiornamento di un URL

Il nostro crawler aggiunge nuovi siti all'Indice ed aggiorna quelli esistenti dopo ogni scansione della rete. In alternativa, è possibile inviarci direttamente l'URL. È importante ricordare che non tutti gli URL segnalati vengono aggiunti all'Indice. Non possiamo quindi fornire indicazioni in merito ai tempi né garantire l'inserimento dell'URL segnalato nell'Indice.

Per segnalare un sito, è necessario inserire l'URL completo, compreso il prefisso `http://`; ad esempio: `http://www.google.com/`. È possibile anche aggiungere commenti o parole chiave che descrivano il contenuto della pagina. Questi elementi vengono utilizzati solo internamente e non influiscono in alcun modo sull'indicizzazione o l'uso del sito da parte di Google.

NOTA: è sufficiente indicare il dominio di primo livello. Non è necessario inviare tutte le singole pagine perché il nostro crawler, Googlebot, è in grado di rilevarle. Nota importante: Google aggiorna l'Indice regolarmente, quindi non è necessario segnalare l'eventuale aggiornamento o cancellazione degli URL. I collegamenti obsoleti vengono rimossi automaticamente alla scansione successiva del crawler ed al momento dell'aggiornamento completo dell'Indice.

URL:

Commenti:

AdSense di Google
Annunci rilevanti al contenuto del vostro sito. Per saperne di più visitate www.google.com/adsense

Annunci su Google senza perdere tempo
[AdWords](#) vi permette di creare annunci altamente mirati e di vederli pubblicati immediatamente su Google.

Posizionamento

Già da prima che i motori di ricerca richiedessero un pagamento per essere presenti nei propri archivi, è diventato evidente che avere un sito identificato da un ricerca, ma in posizione molto bassa, è poco rilevante, in quanto chi consulta i motori di ricerca generalmente non è interessato ai risultati dopo le prime pagine.

I sistemi di posizionamento (ovvero il ranking di un sito) hanno quindi assunto sempre più importanza. Il metodo più semplice per avere un buon ranking è quello di pagare il motore di ricerca. Alcuni motori di ricerca prevedono la possibilità di sponsorizzare i risultati delle ricerche o di essere posti in cima ai risultati attinenti quando viene eseguita una ricerca su determinati termini chiave. Sono così nati termini quali Pay per Inclusion (pagamento per essere inseriti negli indici) e Pay per click (pagamento di una quota ogni volta che un visitatore clicca sul link verso il sito).

Posizionamento e ranking

Ranking

Del concetto di ranking abbiamo già parlato in precedenza dicendo che è l'ordine con cui un motore dispone i risultati della ricerca effettuata. Il ranking maggiore (cioè i primi posti dell'elenco dei risultati) viene assegnato alla pagina che ha maggiore affinità con i termini cercati, a seguire le altre pagine.

Nei primi anni di esistenza, i motori di ricerca assegnavano un ranking maggiore alle pagine in base alla presenza dei termini al loro interno. Ciò si è presto tradotto in pagine con termini ripetuti centinaia di volte al fine di raggiungere il ranking più elevato, ma spesso con contenuti poco vicini a ciò che effettivamente si stava cercando.

Sono quindi stati sviluppati sistemi più elaborati per evitare questi trucchi, che sono considerati spam (tecniche volte ad ingannare il motore di ricerca al fine di apparire nelle prime posizioni) e possono portare ad essere esclusi dagli indici dei motori.

Gradualmente i sistemi usati dai motori di ricerca si sono raffinati, sviluppando tecniche sofisticate in grado di valutare i contenuti delle pagine e a determinare il valore dei link inseriti, basandosi -in estrema sintesi- sul concetto che una pagina che tratta un certo argomento avrà link verso altre pagine con contenuto simile e sarà a sua volta linkata da pagine analoghe.

Google è il motore di ricerca che ha reso massimamente diffuso questo sistema e per la sua precisione e affidabilità si è affermato tra i più popolari e graditi dagli utilizzatori del web.

Page Rank

Il Page Rank (PR) è quindi il punteggio attribuito da Google ad una pagina sulla base di una serie di criteri (in effetti si tratta di un algoritmo matematico complesso e coperto da segreto industriale, peraltro soggetto a continue revisioni al fine di renderlo più efficace e ottenendo così al tempo stesso che nessuno possa creare un sito che rispettando l'algoritmo stesso permetta di conquistare la prima posizione), ma che può essere descritto in maniera semplificata dicendo che non tutti i link hanno lo stesso valore, e che il 'peso' di un link importante ha maggiore valore di un link poco importante.

Il peso di un link dipende a sua volta dal Page Rank della pagina da cui arriva.

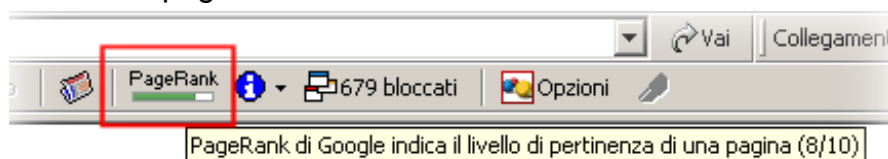
In altre parole una pagina del proprio sito riceverà parte del Page Rank della pagina da cui riceve il link e se la pagina di partenza ha un'alta valorizzazione la pagina del nostro sito ne avrà un beneficio maggiore.

Maggiore è il Page Rank della pagina, migliore sarà il piazzamento che otterrà nella pubblicazione in risposta alle ricerche.

Conoscere il Page Rank delle pagine

Per conoscere il Page Rank delle pagine del proprio sito è sufficiente installare la Google ToolBar messa a disposizione da Google stesso. Si tratta di uno strumento che si integra in Internet Explorer e tramite il quale eseguire direttamente ricerche, ma che presenta anche un indicatore PageRank che mostra il punteggio attribuito alla pagina visualizzata. In effetti il PR mostrato, che varia tra 0 e 10, è un valore approssimato, dato

che Google usa valori di alcuni decimali più preciso, ma è sufficiente per avere una prima indicazione sul valore delle pagine del sito.



Migliorare il posizionamento

Conosciuti il Page Rank delle nostre pagine il passaggio successivo consiste nel migliorarlo.

Migliorare il Page Rank

Sulla base di quanto visto precedentemente, un elemento chiave nella conquista di un page rank migliore è rappresentato da link che arrivano da pagine già ben posizionate.

Questa semplice osservazione ci porta alla considerazione che una pagina deve avere dei buoni contenuti attinenti all'argomento trattato. Meglio ancora essere un'autorità in materia. E' infatti naturale che chi scrive di un certo argomento colleghi la propria pagina a quelle di chi funge da riferimento su quel tema. E' questo il sistema migliore, insieme ad una certa promozione di base, per acquisire importanza e di conseguenza Page Rank.

Tecniche poco corrette sono sconsigliate, se non altro perchè i vari motori di ricerca, da Altavista a Google, i circuiti Fast-AllTheWeb e Inktomi, migliorano costantemente i propri algoritmi di indicizzazione e valutazione, riuscendo anche a riconoscere chi utilizza sistemi di spam. Il risultato di questi comportamenti può forse rivelarsi efficace per breve tempo, ma alla lunga comporta anche il rischio di vedersi esclusi o penalizzati.

Un esempio classico in questo senso è costituito dalla tecnica, ormai superata, di inserire testo con lo stesso colore dello sfondo al fine di inserire termini non visualizzati dagli utenti ma recepiti dai motori di ricerca. Questo sistema viene riconosciuto da diversi motori, che sono in grado di escludere il testo nascosto e di penalizzare la pagina.

Analogamente sistemi di doorway (pagine che permettono di incrementare i link ad un sito) possono causare l'esclusione dell'intero sito dal motore di ricerca.

Ottimizzazione del sito

Realizzare un sito ottimizzato per i motori di ricerca è un'attività che può -e alle volte è preferibile- partire prima della nascita stessa del sito.

Riportiamo qui alcune indicazioni di massima, non necessariamente valide per tutti i motori di ricerca, senza esaurire l'argomento.

Nome del dominio - un sito che riporta già nel suo nome l'argomento di cui tratta è senz'altro attinente al tema in questione

Grafica - i motori di ricerca non sono interessati all'aspetto estetico del sito, anzi un sito molto ricco di grafica probabilmente dà meno importanza al contenuto testuale, che è ciò di cui si preoccupa invece il motore. Siti realizzati con Flash risultano piuttosto indigesti

Struttura della pagina - viene generalmente valutata favorevolmente una pagina semplice, fortemente orientata al testo contenuto che non alla struttura del codice. E'

preferibile evitare la ridondanza di codice (javascript, activeX, ecc) che vada a scapito del messaggio visualizzato

TAG - i tag <title> <description> <keywords> sono stati creati apposta per inserire descrizioni e indicazioni sui contenuti, è quindi opportuno usarli in maniera corretta

Parole chiave (keywords) - sono le parole che dovrebbero consentire di individuare la pagina quando usate per la ricerca. E' opportuno sceglierle con attenzione, dedicando tutto il tempo necessario alla loro scelta. Queste parole devono poi essere presenti nel testo della pagina

Link - la presenza di link all'interno di una pagina è valutata positivamente dai motori di ricerca dato che questa è la strada che usano per valutare l'importanza delle altre pagine.

Conclusioni

Tutte le considerazioni espresse sui motori di ricerca hanno una validità temporale limitata a causa della costante evoluzione delle tecniche adottate. E' quindi opportuno mantenersi costantemente aggiornati in base alle specifiche esigenze.

Non esiste un motore di ricerca migliore degli altri. Può esistere il motore più adatto per certe ricerche o quello che ci risulta più gradevole o pratico nell'utilizzo. E' opportuno verificare costantemente come cambiano nel tempo i vari motori per conoscerli nello stato in cui si trovano effettivamente.

Nel realizzare i siti è opportuno seguire una strategia stabilita a tavolino che può richiedere anche diverso tempo. La realizzazione delle diverse pagine discende poi da questa. Nel creare le pagine è buona norma pensare sempre al fruitore del contenuto e renderlo adatto a lui, non al motore di ricerca. Le norme guida per i motori vengono semplicemente applicate sopra ai contenuti.

Per essere presenti nei motori di ricerca occorre registrarsi perlomeno nei principali motori e directory, solo dopo è il caso di preoccuparsi del proprio livello di rintracciabilità e del Page Rank.

Glossario

Spider: noti anche col nome di Robot, Web Wanderer o Web Crawler, gli Spider sono programmi che si muovono in maniera automatica attraversando la struttura ipertestuale del Web e recuperando in maniera ricorsiva tutti i documenti che sono collegati.

Sono gli strumenti usati dai motori di ricerca per recuperare le pagine da indicizzare.

Metasearch: un particolare tipo di motore di ricerca che interroga altri motori di ricerca e quindi combina i risultati ottenuti. L'utente usa automaticamente più motori con una singola ricerca.



Attribuzione 2.0

Tu sei libero:

- di distribuire, comunicare al pubblico, rappresentare o esporre in pubblico l'opera
- di creare opere derivate
- di utilizzare l'opera per scopi commerciali.

Alle seguenti condizioni:

 **Attribuzione.** Devi riconoscere la paternità dell'opera all'autore originario.

- In occasione di ogni atto di riutilizzo o distribuzione, devi chiarire agli altri i termini della licenza di quest'opera.
- Se ottieni il permesso dal titolare del diritto d'autore, è possibile rinunciare a ciascuna di queste condizioni.

Le tue utilizzazioni libere e gli altri diritti non sono in nessun modo limitati da quanto sopra.

Questo è un riassunto in linguaggio accessibile a tutti della [Licenza completa \(Codice legale\)](#).

[Limitazione di responsabilità](#) 

[Licenza on line - Creative Commons](#)