

## Cercare informazioni in Internet

La cosa che più sconvolge gli utilizzatori di Internet è sicuramente la vastità delle informazioni che vi sono riportate.

Secondo alcune stime, sono attualmente presenti sulla Rete più di un miliardo di pagine contenenti informazioni pubblicamente accessibili. E questo numero cresce con un ritmo esponenziale, visto che negli ultimi due anni è più che triplicato.

Se aggiungiamo a quest'enorme quantità di informazione il fatto che non esiste un metodo standardizzato per l'indicizzazione (come invece avviene ad esempio per i libri, con l'ISBN, codice che identifica univocamente ogni documento stampato), che non esiste un catalogo centralizzato di quanto presente sulla Rete, e che dai documenti presenti sul Web manca anche il nome dell'autore, capiamo bene che ci troviamo di fronte al caos più totale.

### I servizi di ricerca: motori e indici

Immaginate di dover cercare delle informazioni nella più grande biblioteca del mondo, nella quale i libri ed i giornali, privati della copertina e del titolo, sono ammassati sugli scaffali assolutamente a caso, e dove non esiste nessun riferimento ad un catalogo. Un incubo, vero? Questa è un'analogia molto azzeccata della ricerca di informazioni sul World Wide Web, anche se sulla Rete abbiamo l'equivalente di un catalogo, sebbene non sempre aggiornatissimo.

Questo catalogo sono i **servizi di ricerca**, il metodo più sicuro per cercare le informazioni desiderate su Internet. Non pensate comunque che sia facile trovare al primo colpo l'informazione giusta.

Vediamo innanzi tutto quali sono i principali servizi di ricerca, mentre successivamente cercheremo di approfondire il loro funzionamento e imparare alcuni trucchi fondamentali per utilizzarli al meglio e trovare il più rapidamente possibile le informazioni cercate.

Ci sono due tipi principali di servizi di ricerca comunemente utilizzati sul Web:

- i motori di ricerca;
- gli indici tematici.

Per descrivere la differenza tra i due tipi di servizi si può nuovamente ricorrere ad un'analogia libraria. Se pensiamo al Web come se fosse una libreria,

i **motori di ricerca** catalogano tutte le **parole** presenti in ogni libro (sito) presente, e restituiscono l'elenco di tutte le pagine che contengono le parole cercate.

Gli **indici tematici** invece, sono organizzati in base all'argomento del libro ed elencano tutte le copertine (home page) dei libri che appartengono al soggetto cercato.

Per utilizzare un motore di ricerca o un indice tematico, è sufficiente inserire l'indirizzo URL corrispondente nella casella dell'indirizzo del browser e premere **Invio**: verrà visualizzata la **home page** del motore contenente gli elementi utili alla ricerca, come ad esempio la casella di testo in cui inserire le parole da ricercare. Uno dei motori di ricerca più noti è Altavista, il cui indirizzo nella versione italiana è [www.altavista.it](http://www.altavista.it).

È sicuramente opportuno catalogare i principali motori e indici nei preferiti del browser, magari in una apposita cartella, in modo da potervi accedere più velocemente.

### Come funzionano i motori di ricerca

I motori di ricerca consentono di effettuare delle ricerche mediante l'inserimento di una serie di parole chiave, che vengono ricercate all'interno di un enorme archivio. Mediante una serie di criteri, impostati dall'utente o propri del sistema, il motore di ricerca recupera dal suo database i documenti che contengono le parole chiave inserite e presenta i risultati della ricerca, cioè un insieme di indirizzi trovati.

I database utilizzati dai motori di ricerca vengono compilati automaticamente da appositi programmi (detti robot o **spider**) che percorrono autonomamente in lungo ed in largo la Rete per 24 ore al giorno, raccogliendo le informazioni presenti nei siti e seguendo i link presenti in questi siti per raggiungerne altri.

Quando si effettua una ricerca con uno di questi motori, i dati nei quali viene effettuata la ricerca non sono quelli presenti in quel momento su Internet, bensì quelli che sono stati letti tempo addietro dalla Rete mediante i “robot”. Questo sfasamento temporale (che a volte arriva anche a settimane) provoca spesso confusione negli utenti, in quanto può capitare che una pagina indicata come presente ad un certo indirizzo, sia in realtà stata spostata ad un altro, oppure eliminata del tutto. L'estrema dinamicità della Rete è anche questo!

Nessun motore di ricerca cataloga tutto ciò che è presente su Internet, ma solo una minima parte, e ciascuno di essi lo fa in modo differente, per cui l'utilizzo di diversi motori per la stessa ricerca può portare a risultati anche molto differenti.

I fattori che influenzano il risultato sono la *dimensione* del database, la *frequenza* del suo aggiornamento e le *capacità di ricerca*. Altri fattori che differenziano i motori di ricerca sono la loro *velocità*, l'*interfaccia utente*, la *modalità con cui presentano i risultati*, e quanto aiuto offrono ai non esperti mediante degli *help e consigli per la ricerca*.

I motori di ricerca sono ottimi per localizzare una specifica informazione, come ad esempio un documento di cui si conosce una parte del testo o un programma di cui si sa il nome. Lo sono meno per quanto riguarda la ricerca di materiale appartenente ad un dato soggetto.

**I più famosi ed utilizzati motori di ricerca internazionale sono:**

- **Altavista** (<http://www.altavista.com>)
- **Excite** (<http://www.excite.com>)
- **FAST** (<http://www.alltheweb.com>)
- **Google** (<http://www.google.com>)
- **HotBot** (<http://www.hotbot.com>)
- **Infoseek** (<http://www.infoseek.com>)
- **Northern Light** (<http://www.northernlight.com>)



### Spider

indica un programma che navigando nel WWW recupera i documenti. In inglese significa letteralmente “ragno”.

La crescita del numero di motori di ricerca ha dato origine alla creazione di **multi-motori** che, semplicemente, sono dei servizi che **permettono di cercare contemporaneamente le stesse informazioni su più motori di ricerca Internet utilizzando un'unica interfaccia**. Anche se non offrono lo stesso livello di controllo e flessibilità della ricerca dei singoli motori, questi servizi ci permettono con un'unica ricerca di consultare più database, in modo da aumentare la probabilità di trovare le informazioni che cerchiamo, visto che, come abbiamo detto prima, ogni singolo motore di ricerca indicizza solo una piccola parte delle pagine presenti sulla Rete. Questi servizi riordinano i risultati restituiti dai singoli motori di ricerca e li presentano all'utente con modalità differenti, ad esempio ordinati per sito, o per bontà, ovvero per rispondenza alla chiave di ricerca.

**I multimotori più popolari sono:**

- **Dogpile** (<http://www.dogpile.com>)
- **Metacrawler** (<http://www.go2net.com>)
- **ProFusion** (<http://www.profusion.com>)
- **SavvySearch** (<http://www.savvysarch.com>)

### Come funzionano gli indici tematici

Gli indici tematici sono **organizzati gerarchicamente in categorie** che consentono all'utente della Rete in cerca di informazioni di scorrere degli **elenchi di siti Web suddivisi per soggetto**, un po' come se si consultassero le pagine gialle. Questi elenchi sono compilati e aggiornati da **personale “umano”** (a differenza dei “robot” di prima), e sono generalmente **strutturati sotto forma di “alberi”**, composti da poche categorie principali (le radici) che si diramano in sotto categorie, sotto sotto categorie, ecc., fino ad arrivare agli elementi veri e propri (le foglie) contenenti le informazioni. Gli **archivi di questi indici tematici sono più piccoli rispetto a quelli dei motori di ricerca**, quindi i risultati delle nostre ricerche tendono ad essere più concisi.

Un'altra differenza è che, mentre un motore di ricerca normalmente indicizza ogni pagina di un sito web, gli indici tematici **contengono solo un riferimento alla home page del sito**. Inoltre, dato che questi indici sono aggiornati mediante l'intervento umano, è molto ridotta la possibilità di essere rimandati a siti che non hanno nulla a che vedere con quello che stiamo cercando.

Gli indici tematici sono sicuramente il modo migliore per ricercare le informazioni partendo da un soggetto generico, invece che da un dato ben specifico.

Alcuni dei più famosi indici tematici con dominio italiano sono:

- **Arianna** (<http://www.arianna.it>)
- **Yahoo** (<http://www.yahoo.it>)
- **Lycos** (<http://www.lycos.it>)
- **Virgilio** (<http://www.virgilio.it>)

## Metodi di ricerca

Cerchiamo ora di capire alcune strategie e modalità di ricerca più efficaci per evitare di finire nei siti più incredibili, che non hanno nulla a che fare con quanto stiamo realmente cercando.

Gli indici possono essere un ottimo punto di inizio della ricerca se il soggetto che cerchiamo è presente come elemento di questo indice.

Gli indici sono anche ottimi per trovare informazioni su di un argomento quando non abbiamo le idee ben chiare sui dettagli della nostra ricerca. Anche gli indici tematici forniscono normalmente la possibilità di inserire una parola chiave, con la differenza che questa non viene ricercata all'interno delle pagine Web, ma solo utilizzata per trovare più rapidamente le categorie che ci interessano, senza dover percorrere categorie e sotto categorie.

Per ogni elemento viene indicato il percorso per arrivare all'argomento attinente alla parola richiesta.

Questi **indici** coprono solo una piccola parte delle informazioni contenute sul Web, e quindi sono utili principalmente per la ricerca di informazioni su soggetti di interesse comune o scolastico. Se occorre cercare qualcosa di specifico, dobbiamo rivolgerci ai motori di ricerca.

L'utilizzo di un motore prevede l'inserimento delle parole da ricercare in una apposita casella di testo. Le parole inserite verranno cercate nel database del motore di ricerca e gli indirizzi di tutte le pagine che le contengono verranno presentati all'utente; nell'elenco verranno indicate per prime le pagine che più rispondono ai criteri richiesti.

All'inizio dell'elenco viene indicato il **numero di pagine trovate**, e per ciascuna pagina vengono fornite indicazioni relativamente a:

- il titolo della pagina;
- alcune righe della pagina che dovrebbero fornire un'idea del contenuto;
- l'indirizzo URL della pagina;
- la data in cui la pagina è stata modificata per l'ultima volta (serve per capire se la pagina è aggiornata o no);
- la lingua in cui è scritta la pagina.

Spesso tuttavia, se le parole inserite sono poco specifiche, vengono restituite migliaia di pagine come risultato. Per utilizzare un motore di ricerca con successo, è quindi essenziale saper applicare delle tecniche che riducano il numero di pagine restituite, e che mettano in evidenza le pagine più rilevanti per la nostra ricerca.

Vediamo alcune possibilità.

### 1) Inserire più parole separate da spazi.

Occorre scomporre l'argomento in alcuni concetti chiave e inserire tutte le parole individuate come chiavi di ricerca.

Alcuni motori restituiscono tutte le pagine che contengono almeno una delle parole richieste e quindi anche quelle che ne contengono solo una e che magari non riguardano l'argomento di nostro interesse (si dice che il motore utilizza come operatore standard l'**OR**).

Altri motori restituiscono le pagine che contengono tutte le parole inserite (il motore utilizza come operatore standard l'**AND**).

**2) Inserire un'intera frase racchiusa fra virgolette**, se si vuole che le pagine trovate contengano esattamente la frase e non le singole parole.



Utilizzare le virgolette può essere molto utile nei seguenti casi:

- cerchiamo un preciso documento di cui conosciamo almeno in parte il contenuto: per una ricerca più sicura e precisa conviene inserire una frase contenuta nel testo del documento;
- ricerchiamo un personaggio: ci conviene inserire il nome e cognome del personaggio racchiuso fra apici per evitare di trovare anche tutte le pagine che fanno riferimento a personaggi con lo stesso nome o lo stesso cognome;
- cerchiamo termini come ad esempio “borsa di studio” o “fondo di investimento”, formati da parole che prese singolarmente hanno un significato diverso e che quindi vanno considerate come fossero un elemento unico.

### 3) Utilizzare i simboli + e - prima delle parole.

L'utilizzo del + o del - prima di una parola o di una frase serve per dare indicazioni più precise su di essa nel seguente modo:

- facendo precedere la parola dal simbolo + (attenzione a **non inserire spazi in mezzo**), si vuole indicare che la parola deve necessariamente essere presente nella pagina;
- facendo precedere la parola dal simbolo - si vuole indicare che la parola non deve essere presente nella pagina (da utilizzarsi per restringere i risultati).

È chiaro che utilizzando più parole opportunamente combinate con + e - possiamo rendere più mirata la nostra ricerca.

Se vogliamo informazioni relative ai musei di Parigi escluso il Louvre possiamo impostare la ricerca così:  
+musei +Parigi -Louvre

**4) Inserire una parte di parola seguita da \*** (carattere jolly che sostituisce una o più lettere): vengono ricercate tutte le pagine contenenti parole che iniziano con la parte di parola specificata. L'asterisco può quindi essere usato per cercare tutte le declinazioni di una parola, come il genere ed il numero dei sostantivi o il tempo e il modo dei verbi. L'asterisco può essere usato anche all'interno di una parola, ad esempio quando non si è sicuri della sua ortografia. Bisogna ovviamente non eccedere nell'utilizzo dell'\* perché in alcuni casi potremmo avere una quantità eccessiva di risultati.

“cavall\*” sta per cavalla, cavallo, cavalli, cavallerizzo, e per tutte le altre parole che cominciano per “caval”, “vend\*” sta per venderò, venduto, vendendo, vendesi, ecc

### Ricerca avanzata

Alcune altre funzioni fanno parte della cosiddetta ricerca avanzata (normalmente occorre selezionare un apposito pulsante nella home page del motore per passare ad una pagina specifica per ricerca avanzata).

In figura potete vedere la finestra utilizzata dal motore Altavista.it per la ricerca avanzata.

La ricerca avanzata consiste essenzialmente nell'utilizzo degli **operatori logici** (o operatori booleani) **AND**, **OR**, **NOT** e **NEAR** che servono per creare delle relazioni tra parole specificate per la ricerca:

- l'operatore **AND** posto fra due parole indica che entrambe le parole devono essere presenti nella pagina;
- l'operatore **OR** posto fra due parole indica che anche solo una delle due parole può essere presente nella pagina;
- l'operatore **NOT** posto prima di una parola (separato da uno spazio e preceduto dall'operatore **AND**) indica che la parola non deve essere presente nella pagina;
- l'operatore **NEAR** posto fra due parole indica che le parole devono essere entrambe presenti, ma ad una distanza massima di 10 parole l'una dall'altra (è più restrittivo dell'operatore **AND**, ma meno che unendo le due parole in una frase).

Con la stringa: “Romanticismo AND Leopardi AND NOT Manzoni”, vogliamo trovare informazioni sul Romanticismo e Leopardi escludendo Manzoni.

L'operatore **OR** può essere utilizzato anche per i sinonimi: quando una parola che ricerchiamo ha dei sinonimi o forme diverse (anche solo singolare e plurale) è opportuno inserire nella ricerca le diverse parole equivalenti, unite appunto mediante l'operatore **OR**.

Per restringere il più possibile il numero di risultati si può usare l'operatore **OR** in combinazione con l'operatore **AND**, ma in questo caso è necessario utilizzare le parentesi, cioè racchiudere le parole unite dall'operatore **OR** tra parentesi.

La ricerca avanzata di alcuni motori (ad esempio Altavista) permette inoltre di ordinare i risultati e di specificare nell'interrogazione un intervallo di date per avere solo le informazioni pubblicate in quel periodo.



Riassumendo, proviamo ad elencare i passi da seguire per **organizzare una ricerca avanzata**.

- 1) Con una o due frasi descrivere quello che si vuole cercare sulla rete ed eventualmente quello che si vuole escludere. Esempio: *Quando migrano le rondini?*
- 2) Sottolineare i concetti base della ricerca, ovvero le chiavi di ricerca. Esempio: Quando migrano le rondini.
- 3) Elencare i sinonimi e le varianti delle parole da cercare. Esempio: migrano, migrazione, migrazioni, rondine, rondini.
- 4) Combinare le parole chiave, i sinonimi e le varianti con gli operatori logici. Utilizzare l'operatore OR per combinare i sinonimi e le varianti, racchiudendoli tra parentesi, e utilizzare AND per combinare le parole, o le espressioni, chiave. Esempio: (migrano OR migrazione OR migrazioni) AND (rondine OR rondini).
- 5) Controllare l'ortografia. Se scrivete una parola sbagliata, il motore di ricerca troverà solo le pagine dove questa parola è scritta sbagliata, sempre che ne esistano!

#### L'operatore OR

Il comportamento predefinito di **Google** è quello di considerare tutte le parole in una ricerca. Se desideri cercare specificamente *una parola o un'altra* tra vari termini, puoi utilizzare l'operatore OR (nota che devi digitare **OR** tutto in MAIUSCOLO). Ad esempio, [ San Francisco Giants 2004 OR 2005 ] restituirà risultati su uno di questi anni, mentre [ San Francisco Giants 2004 2005 ] (senza OR) mostrerà le pagine che contengono entrambi gli anni nella stessa pagina. L'operatore OR può essere sostituito dal simbolo |. **L'operatore AND invece rappresenta l'impostazione predefinita, quindi non è necessario aggiungerlo.**

### Approfondimento: La Ricerca

Ci sono metodi più sofisticati basati sulla ricerca della parola non all'interno del testo della pagina, ma negli elementi delle pagine come il **titolo**, il **dominio**, L'**URL**, i **link**. L'efficacia di una ricerca aumenta se si combinano questi metodi particolari con quelli visti prima.

Supponiamo per esempio di voler cercare informazioni sull'infanzia di Alessandro Manzoni. È molto probabile che esistano pagine Web a lui dedicate, nel cui titolo appaiano proprio il nome e cognome del Manzoni. Potremmo quindi provare a ricercare la seguente stringa. **title:"Alessandro Manzoni" AND infanzia**. Questa stringa richiede al motore di ricerca di restituire tutte le pagine Web nel cui titolo appare la frase "Alessandro Manzoni" e nel cui testo appare la parola "infanzia". È fondamentale che dopo l'operatore "title:" non appaia nessuno spazio.

Oltre alla ricerca per titolo, altri campi utilizzabili per restringere i risultati della ricerca sono,

- **domain** per prendere in considerazione solo i domini contenenti la parola, specificata; ad esempio domain:.it per avere solo domini italiani
- **URL** per avere le pagine il cui URL contiene la parola specificata; ad esempio,URL:Torino se vogliamo tutte le pagine i cui indirizzi contengono la parola "Torino",
- **link** per sapere quali sono i siti che contengono al loro interno dei link ad un certo sito.

### Qualche consiglio finale

Non sempre effettuare subito ricerche sofisticate con l'inserimento di molti termini e l'utilizzo di tecniche avanzate è vantaggioso: potrebbe infatti eliminare dai risultati anche dei dati interessanti.

Conviene perciò iniziare con una ricerca tradizionale e semplice, e solo se i risultati ottenuti sono troppo numerosi e poco confacenti alle nostre necessità, allora raffinare la ricerca utilizzando tecniche avanzate.

Ogni motore di ricerca può disporre di funzionalità particolari e soprattutto, in una realtà come quella di Internet in continua evoluzione, **queste possono cambiare**. Le indicazioni date qui sono di tipo generale e vanno adattate alla situazione trovata nella realtà del momento.

Consigliamo pertanto di consultare la guida di cui ogni motore dispone e nella quale sono spiegate, quasi sempre in modo molto semplice e chiaro (lingua inglese a parte se utilizziamo un motore non italiano) le modalità di ricerca da utilizzare.

Quando apriamo una pagina fra quelle ottenute come risultato di una ricerca, conviene sempre risalire alla home page del sito di cui la pagina fa parte (normalmente esistono nella pagina link specifici per questo), per renderci conto di che tipo di sito si tratta, da chi è stato realizzato e con quali obiettivi, informazioni utili per poterne meglio valutare il contenuto.

Sotto è riportata una parte della guida del motore Altavista (versione italiana).

### RICERCA RISULTATO

monna lisa Trova i documenti che contengono sia monna che lisa e tutte le combinazioni di maiuscole (Monna, MONNA, liSA, Lisa). AltaVista ti mostra prima i documenti che contengono entrambe le parole, raggruppandoli e mettendoli all'inizio del documento.

Monna Lisa TROVA SOLO - Trova i documenti che contengono sia Monna che Lisa ma non quelli che hanno combinazioni di maiuscole. Quando usi una parola con la maiuscola, AltaVista considera che tu stia cercando solo una corrispondenza esatta.

+monna +lisa Trova solo i documenti che contengono entrambe le parole.

“Monna Lisa” TROVA SOLO - Trova i documenti che contengono entrambe le parole, con la maiuscola, disposte una di fianco all'altra. Se delimiti con le virgolette una serie di parole, queste vengono interpretate come una frase; in pratica dici a AltaVista che sei interessato solo a quei documenti che contengono queste parole in questo ordine.

+ “Monna Lisa”-Louvre Trova i documenti che contengono la frase “Monna Lisa” e non contengono la parola “Louvre”

+Monna +Lis\* TROVA SOLO - Trova i documenti che contengono Monna e tutte le parole che cominciano con Lis. Ti consigliamo di usare questa funzionalità quando non sei sicuro di come è scritta la parola.

## Approfondimento: Elenchi di Motori

Per mantenersi aggiornati sui motori disponibili e sul loro utilizzo, esistono siti appositi che contengono elenchi di motori. Fra questi ricordiamo:

- [www.searchenginewatch.com](http://www.searchenginewatch.com) il più famoso a livello internazionale
- [www.motoridiricerca.it](http://www.motoridiricerca.it) il più famoso in Italia

## I Portali:

**Portal o web portal:** un sito www che facilita l'ingresso di un utente nella rete guidando verso altri siti di particolare interesse.

È un servizio alternativo ai motori di ricerca.

Organizza le informazioni in categorie e consente all'utente di ritrovare e caricare velocemente le pagine alle quali l'utente vuole ritornare.

## Esercizi

### 1) Accedere ad una pagina di cui si conosca l'indirizzo.

- a) Accedi al sito del quotidiano Repubblica ([www.repubblica.it](http://www.repubblica.it)) e naviga attraverso le notizie del giorno.
- b) Accedi poi al sito del quotidiano La Stampa ([www.lastampa.it](http://www.lastampa.it)) e rifletti sul diverso modo di presentare le notizie.

### 2) Aggiungere ai preferiti.

- a) Crea nei Preferiti una cartella Quotidiani.
- b) Accedi al sito del quotidiano Repubblica e catalogalo come preferito nella cartella Quotidiani.
- c) Fai la stessa cosa per il sito del quotidiano La Stampa.
- d) Accedi nuovamente al sito della Repubblica selezionandolo dai preferiti.

### 3) Navigare all'interno di un sito.

- a) Accedi al sito dell'editore Petrini ([www.petrini.it](http://www.petrini.it))
- b) Naviga attraverso i link per cercare il tuo libro di testo.

### 4) Ricercare informazione.

- a) Ricerca informazioni relative alla letteratura italiana dei '900 utilizzando un indice tematico (ad esempio [www.yahoo.it](http://www.yahoo.it)).
- b) Ricerca le stesse informazioni utilizzando un motore di ricerca (ad esempio [www.altavista.it](http://www.altavista.it)).
- c) Confronta i risultati ottenuti.
- d) Salva in un file una pagina che ritieni significativa

### 1) Ricercare informazioni su un personaggio.

- a) Ricerca informazioni relative al tuo cantante o gruppo preferito.
- b) Salva un'immagine come sfondo dello schermo
- c) Salva un'altra immagine in un file.
- d) Stampa una parte di testo che ti interessa.