

COSA SONO I PROTOCOLLI

Miliardi di bit viaggiano ogni giorno sulla **Rete**. Vi siete mai chiesti **come fanno ad arrivare al corretto destinatario?**

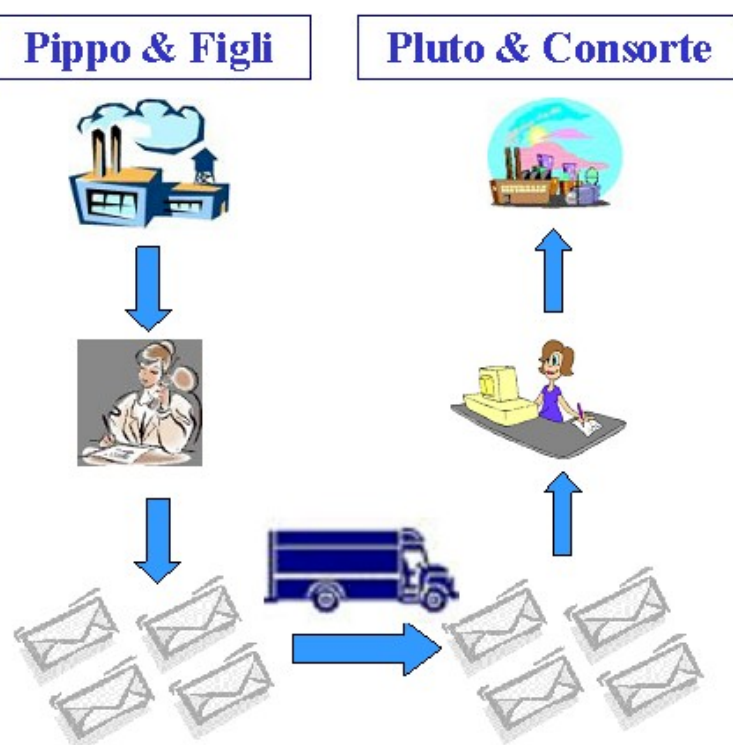
I **protocolli** sono la risposta a questa domanda!

I **protocolli** sono le **regole utilizzate** per la **trasmissione di qualsiasi tipo di dato su Internet** (compresa la voce). Già, ma allora **che cosa è un protocollo** ? Un protocollo è una **descrizione formale** del **formato dei messaggi** e delle **regole che devono seguire due computer** affinché lo **scambio** (di bit) possa avvenire.

Un **protocollo** può contenere **regole estremamente dettagliate**, come quelle che **identificano il significato** di ogni **singolo bit** nella **costruzione di un messaggio**, oppure fornire uno **scenario di alto livello**, come per esempio definire come avviene il **trasferimento di un file da un computer a un altro**. I **protocolli** possono essere di **basso livello** (come **spedire i bit e i byte** attraverso la rete) o di **alto livello** (come **scambiarsi documenti, file o altri oggetti**).

Per esempio, il **TCP/IP** (Transfer Control Protocol) è un protocollo di **basso livello**, l'**HTTP** (Hyper Text Transfer Protocol) è un **protocollo di alto livello**.

Ma vediamo un **esempio concreto** che descrive **cosa si intende per protocollo**:



Esempio: Il **direttore** della **Pippo e Figli** manda una **lettera riservata** al **direttore** della **Pluto e Consorte**.

Primo protocollo.

Il modo con cui i **due comunicano**, per esempio i **riferimenti a lettere precedenti**, lo **stile della lettera**, il **modo di salutare alla fine della lettera**, e così via, rappresenta il **protocollo ad alto livello**, cioè quello **applicativo**.

Per spedire la lettera il **direttore lo passa alla sua segretaria**. Ciò avviene secondo le **regole interne** della **Pippo e figli**, ed è perciò un'**interfaccia**.

Secondo protocollo.

La **segretaria prende la lettera**, la mette in una **busta** aggiungendo il **nome del destinatario** e la scritta **RISERVATO**.

Queste **informazioni** sono per la sua **controparte** nella **Pluto e consorte**, ed è quindi **anch'esso un protocollo**.

Terzo protocollo.

La **busta** viene quindi passata all'**Ufficio Posta dell'edificio** secondo la **procedura ordinaria (altra interfaccia)**, il quale aggiunge l'**indirizzo completo**, il **CAP** e **altre informazioni** necessarie alla **spedizione**, e la passa quindi al **corriere**, che rappresenta il **meccanismo fisico di trasferimento del messaggio**.

Quando la **lettera arriva**, questa viene gestita dall'**Ufficio Posta della Pluto e consorte** che, **dopo aver buttato la busta esterna** con l'indirizzo dell'edificio, **la passa alla segreteria del direttore**.

Questa registra l'arrivo della missiva, la **toglie dalla busta più interna**, e poi **consegna la lettera vera e propria** al **direttore della Pluto e consorte**.

È evidente che **perché il sistema funzioni** bisogna che i **vari protocolli** siano **rispettati da entrambe le aziende**.

Mentre infatti **le interfacce** possono essere **diverse (ogni azienda ha le sue procedure interne)**, i **protocolli devono essere stati concordati prima**, altrimenti alcune informazioni potrebbero andare perse, o addirittura la lettera potrebbe non essere recapitata in tempo.

Pensate a una **segretaria italiana** che **riceve una busta** da una **consociata asiatica** con sopra scritto **URGENTE in cinese!** Ne nasce una considerazione importante.

La **base di ogni protocollo** è il **concetto di standardizzazione**.

Più vasta è l'**accettazione dello standard**, più **forte e diffuso** è il **protocollo**. Gli **standard internazionali** sono in genere i più importanti, ma non sempre.

Un esempio è proprio il **TCP/IP**, nato per volontà dell'**agenzia americana DARPA** (Defense Advanced Research Projects Agency) e poi diventato **di fatto il maggior sistema di protocolli** per l'**interconnessione di reti a livello mondiale**.

SERVER

Naturalmente un **Server Internet** deve essere **connesso alla Rete 24 ore su 24 per 365 giorni l'anno** quindi solitamente fa parte di una **LAN (Local Area Network)** connessa ad Internet attraverso un Provider di accesso (ISP) oppure direttamente ad una compagnia telefonica.

Il proprietario di un Server Internet può essere a sua volta un Provider di presenza **hosting provider** (fornitore di spazio web) oppure una azienda o un privato.

La **tipologia di Server Internet** più diffusa è quella di **Server HTTP o Web Server**, ma come vedremo tra poco non è l'unica!

Con il termine "**Web Server**", più specificatamente, ci si riferisce al **programma "Server"** che ha il compito di gestire le "**pagine Web**" mediante il protocollo **HTTP** (Hyper Text Transfer Protocol) .Un computer **Server Internet** può "ospitare" ed "eseguire" più di un programma "Server" contemporaneamente: un **Web Server**, un **FTP Server**, un **Mail Server**, un **News Server**, un **POP Server**, ecc., anche se di solito è dedicato ad uno o pochi compiti specifici.

Lo schema per esempio mostra il funzionamento di un **Web Server**.



Di seguito sono elencati i **più comuni server Internet** commerciali:

- Server Web (o HTTP)
- Server FTP
- Server e-mail (POP3 per la posta in entrata e SMTP per la posta in uscita)
- Server News
- Server Gopher

- Server di login remoto (TELNET)

Server http

I server HTTP sono divenuti sempre più popolari e spesso sono il sinonimo di ciò che molte aziende intendono quando sentono parlare di server Internet.

Tuttavia, sebbene il Web sia uno strumento complessivo considerevole, la funzionalità di un Server Internet non si misura soltanto in base a questo servizio.

Server FTP

Un server FTP è un computer dedicato alla distribuzione di files agli utenti: in pratica funziona come un magazzino o repository di files di qualunque tipo suddivisi in cartelle e sottocartelle esattamente come in un qualsiasi disco fisso di windows.

Server e-mail

E' un "**ufficio postale**" **virtuale**: si tratta di uno o più computer dedicati allo smistamento di messaggi di posta elettronica e alla gestione delle caselle di posta corrispondenti agli indirizzi e-mail.

Ogni casella di posta è in pratica una porzione di disco fisso con 2 precisi limiti assegnati, il primo è relativo alla dimensione massima espressa in multipli del Byte (KiloBytes, MegaBytes, GigaBytes, ecc.) che può contenere nel suo insieme ed il secondo è la dimensione massima di ogni messaggio (compresi gli allegati)

Il **Server POP3** è dedicato alla **posta in entrata**, mentre il **Server SMTP** (Simple Mail Transfer Protocol) è dedicato alla **posta in uscita**.

Server News

L'azienda vuole avere l'accesso ai newsgroup UseNET? Un server news potrebbe essere sfruttato per ospitare i propri newsgroup su Internet o all'interno della società stessa.

UseNet è il nome di un formato newsgroup che si è evoluto in un'ottima fonte di notizie aggiornate su argomenti specifici.

Se un newsgroup non esiste, chiunque può crearlo: esso potrebbe riferirsi a un particolare argomento oppure fornire informazioni di supporto prodotto oppure raccogliere informazioni per il feedback.

Attualmente esistono migliaia di newsgroup e tentare di monitorare quelli che potrebbero avere un impatto sull'azienda potrebbe rivelarsi un'ottima strategia.

Creando un newsgroup diventa possibile offrire ai clienti o ai potenziali clienti tutte le informazioni che essi possono desiderare conoscere e l'occasione per poter discutere eventuali reclami o per poter accogliere suggerimenti o idee innovative.

I newsgroup UseNet generalmente sono considerati una sorta di avviamento da parte delle aziende che li implementano e li sfruttano.

Server Gopher

Nonostante il rapido sorpasso del Web ai danni di altri servizi, tra cui i server Gopher, questi ultimi non sono da riporre nel dimenticatoio.

Se l'azienda desidera creare un **catalogo di informazioni organizzato in menu** per poter accedere alle numerosissime risorse di Internet, un **server Gopher** potrebbe essere la soluzione ideale.

Gopher è stato il primo sistema informativo che ha facilitato la navigazione in Internet. I client Gopher visualizzano un **menu di opzioni sotto forma di testo** che l'utente può selezionare per portarsi nella posizione desiderata.

Spesso, prima di arrivare a destinazione, è necessario effettuare selezioni in più menu, tuttavia non è difficile amministrare una simile architettura se viene organizzata bene fin dall'inizio.

I nomi dei file e delle sottodirectory equivalgono alle voci di menu del client Gopher: l'utente deve limitarsi a fare clic sulla voce appropriata e l'argomento verrà visualizzato sullo schermo.

Server di login remoto

Alcune reti necessitano tuttora di un login remoto (per esempio Telnet) per consentire l'accesso alle risorse sul server. La Microsoft sembra minimizzare l'utilità di un server Telnet su Windows NT (non esistono infatti server Telnet NT, ma solo software esterni), ma in presenza di un'esigenza o di una preferenza specifica, questa soluzione è l'ideale per un login remoto dalla riga di comando.

Telnet consente all'utente di collegarsi a un computer remoto e di sfruttare i servizi offerti dallo stesso sito. Gli utenti che devono viaggiare spesso fino a non molto tempo fa solitamente dovevano affrontare i costi tutt'altro che bassi delle telefonate interurbane dalla località dove si trovavano fino alla sede del loro ufficio

per potersi collegare via modem alla rete locale e inserire informazioni in un modulo dati o prelevarne altre da un database.

A mano a mano che Internet è entrata nelle aziende, ai viaggiatori è stata offerta la possibilità di connettersi, ovunque si trovino, alla Rete attraverso un provider locale e di usare Telnet per ottenere lo stesso accesso ai dati che avevano prima.

Questa possibilità per esempio permette a un venditore di connettersi alla rete anche mentre è in viaggio e di registrare gli ordini in un'applicazione che è la stessa che viene usata dai venditori in sede.

Motivazioni ed esigenze aziendali per l'installazione di un Server internet

Prima di fare una valutazione delle **motivazioni** per cui **la presenza di un'azienda su Internet** potrebbe portare molti benefici, è **essenziale porsi alcune domande** relative alle principali esigenze specifiche dell'azienda.

Innanzitutto è necessario determinare lo **scopo del server Internet** che si vuole installare, quindi quale sarà la sua **funzione primaria** in qualità di **punto centrale di presenza su Internet**. Di seguito sono proposte alcune domande:

- Che tipo di connessione a Internet usare? L'azienda ha bisogno di una connessione T1 ad alta velocità per un'**ampiezza di banda** considerevole? Oppure il server servirà per connessioni dial-in? La connessione a Internet deve essere decisa sulla base dei costi, della disponibilità di servizi e delle esigenze specifiche dell'azienda.
- Che cosa si vuole ottenere dalla presenza su Internet? L'azienda necessita dell'installazione di servizi e-mail per gestire le comunicazioni con i clienti e i dipendenti? È interessata a una connessione ad alta velocità per sfruttare un sito Web come opzione pubblicitaria? Le esigenze e la priorità evidenziati in ciascuna domanda sono elementi fondamentali da mettere a fuoco.
- Quali sono gli altri benefici che potrebbero derivare da Internet? Oltre alle esigenze primarie di networking, ci sono altri benefici quali per esempio un facile accesso a materiale di ricerca o ai rivenditori che potrebbero aumentare le ragioni valide per connettersi a Internet?
- Quali sono i benefici tangibili che si possono ottenere? I profitti dell'azienda che beneficio possono trarre dalla presenza su Internet? Attraverso la

pubblicità, il marketing o una migliore assistenza post-vendita sarà possibile realizzare profitti maggiori? Ci sarà la possibilità di risparmiare?

- Quali sono i benefici intangibili che si possono ottenere? La presenza su Internet potrebbe migliorare le relazioni con i clienti, l'immagine dell'azienda o la credibilità? I benefici intangibili sono difficili da individuare ma in effetti sono quelli che pesano di più sulla bilancia della decisione.

SITO WEB

La parola "**sito**" in origine si riferiva ad un'**istituzione (edificio, ufficio, organizzazione) contenente un computer** con le funzioni di **server in rete**.

Oggi, invece, secondo la terminologia più accettata pubblicamente, la parola "**sito**" si riferisce anche ad una **raccolta organizzata di file (pagine Web o d'altro tipo) accessibile tramite uno stesso nome di dominio**.

I **Siti Web** sono costituiti da **una o più pagine** che gli utenti visitano durante l'esplorazione del **World Wide Web**.

La **home page** è la **prima pagina del sito Web** e costituisce la "**copertina**" o l'**indice** che consente di accedere alle informazioni disponibili all'interno del sito. La prima pagina di un sito si chiama home page **anche se il sito contiene una sola pagina**.

Le **pagine del World Wide Web** sono **diverse dalle pagine di un libro**. Ad esempio, una pagina Web **non ha dimensioni predefinite** e può essere molto lunga o molto corta. La stampa di una **pagina Web** può infatti richiedere diversi fogli. Anziché da numeri di pagina, ogni pagina Web è identificata da un indirizzo, o **URL**.

Una **pagina Web** può essere molto semplice e contenere solo poche righe di testo, oppure molto elaborata, piena di colori, di tipi di caratteri diversi e immagini, magari animate o che cambiano forma al passaggio del mouse.

Una **pagina Web** può anche contenere **suoni** e altri **effetti multimediali**. Alcune pagine sono talmente belle che sembrano realizzate da uno studio cinematografico.

Oltre che belle da vedere, le pagine Web hanno una caratteristica unica: i collegamenti ipertestuali. Un collegamento ipertestuale è un riferimento a un'altra pagina o a un'altra posizione del Web e può essere costituito da un testo o da un'immagine.

Quando si fa clic con il mouse sul collegamento, il programma client si collega a un altro sito Internet, che potrebbe trovarsi dall'altra parte del mondo. I collegamenti ipertestuali consentono quindi di passare rapidamente da un sito all'altro.

I collegamenti ipertestuali sono uno strumento impagabile. Immaginate ad esempio

di andare nella biblioteca della vostra città per svolgere una ricerca su di un argomento scientifico. Dopo aver recuperato un testo e aver preso i primi appunti, vi accorgete che, per approfondire l'argomento, dovete consultare un altro libro. Il secondo libro porta quindi a un terzo, che a sua volta fa riferimento a un quarto e così via. I collegamenti ipertestuali consentono di evitare gran parte di questo noioso procedimento. Quando l'utente fa clic su un collegamento ipertestuale, il browser visualizza in pochi secondi il contenuto del nuovo sito.

Se la nuova pagina contiene collegamenti ad altri siti, è sufficiente fare un altro clic per visualizzare tutto il materiale desiderato.

Il World Wide Web (WWW)

Il World Wide Web, o più semplicemente "il Web", è un servizio Internet di grandi dimensioni e, dal 1993, il più conosciuto dei servizi disponibili. In effetti molte persone pensano che Internet e il WWW siano sinonimi e alcuni non si rendono conto che il Web è in realtà un'entità che funziona all'interno di Inter(Wide net.

La maggior parte di voi troverà che il Web è il modo più comodo per avere una vostra presenza in Internet.

Il Web contiene un numero praticamente infinito di documenti, che vengono presentati utilizzando una vasta gamma di mezzi di comunicazione, dai semplici documenti testuali a quelli con caratteristiche multimediali.

La caratteristica veramente unica dei **documenti WWW** è che essi sono collegati ad altri documenti per mezzo di una tecnologia nota come **ipertesto**.

L'ipertesto vi permette di passare da un documento ad un altro relativo allo stesso argomento, con un semplice clic del mouse selezionando una parola o una immagine, che siano state impostate come link o collegamento ipertestuale.

Un "link" è normalmente identificato da una immagine o da una parola evidenziata, rispetto al testo normale, mediante l'utilizzo di un colore differente o del carattere grassetto. Di solito il documento "legato" che viene richiamato a video,

definisce o elabora maggiormente il termine o l'elemento grafico utilizzato come link, nel documento di partenza.

Ad esempio, un catalogo potrebbe avere dei link ipertestuali con gli elenchi di specifiche dei prodotti, consentendo ai potenziali clienti di accedere in modo semplice a informazioni più dettagliate, riguardanti i prodotti di loro interesse. L'ipertesto rende il Web veramente interattivo.

Quando si sfogliano le risorse Web, siete voi a decidere come navigare attraverso uno specifico documento. In altre parole non è necessario che leggete una risorsa Web nel modo lineare in cui viene letta la maggior parte dei libri. Potete saltare da un punto all'altro del documento, in base ai vostri specifici interessi. Il Web supporta inoltre altri protocolli quali FTP, Gopher, e-mail e NNTP.

Il progetto **WWW** è stato sviluppato presso il **CERN** (centro europeo per la fisica nucleare) in **Svizzera**. Il protocollo **HTTP Hypertext Transfer Protocol** costituisce il **fondamento del Web**.

In origine era inteso come mezzo grazie al quale i fisici erano in grado di condividere le reciproche informazioni, ma ben presto il **protocollo HTTP** venne incorporato in Internet, quando si comprese quanto fosse vasto il campo di applicazione della sua tecnologia.

Farsi il sito

Per fare un uso efficace delle tecnologie Internet, dovete pensare al modo migliore di utilizzare la tecnologia e Internet.

Dovete tenere conto dei vostri reali obiettivi, del vostro pubblico potenziale, delle vostre motivazioni nell'utilizzo di una determinata tecnologia e della quantità di tempo e denaro che volete dedicare al progetto. Se siete un utente dilettante di computer, forse desiderate costruirvi un sito Web soltanto per divertimento e perché si tratta di una sfida.

Se avete una vostra attività commerciale, potreste avere bisogno di una presenza in Internet per migliorare il servizio da offrire ai vostri clienti o l'esigenza di impostare una intranet aziendale per migliorare la comunicazione all'interno della vostra organizzazione.

O, magari, la vostra intenzione è quella di avviare un'attività commerciale basata esclusivamente su Internet e fornire servizi collegati a Internet.

Qualunque sia la natura delle vostre motivazioni, dovete sviluppare un vostro business plan e decidere quali servizi e tecnologie è appropriato per voi utilizzare.

Che cosa significa essere presenti in Internet?

In termini molto semplici, il fatto di avere una propria presenza in Internet significa avere la possibilità di far conoscere e di rendere disponibili in Internet le vostre idee e i servizi che siete in grado di offrire.

Naturalmente il modo di procedere dipende dai vostri interessi e dai vostri obiettivi finali e, solitamente, varia a seconda degli interessi individuali di chi imposta un nuovo sito.

È necessario tenere conto anche di altre questioni di ordine pratico: non basta pensare a quello che il vostro sito Internet ideale dovrebbe comprendere, bisogna anche pensare a quanto tempo e denaro potrete investire, non soltanto per l'impostazione del sito, ma anche per la sua manutenzione.

Per una presenza Internet può essere necessario investire risorse considerevoli, ma i vantaggi che derivano dallo sviluppo e dalla manutenzione di un proprio sito, sono notevolmente maggiori del fatto di pubblicare i propri contenuti in uno spazio "affittato" presso un server.

Dal punto di vista della manutenzione, avere la responsabilità del vostro server garantirà il completo controllo del vostro sito. Se dovrete effettuare modifiche o aggiornamenti al vostro sito o se avrete qualche difficoltà con l'hardware, potrete cambiare o riparare immediatamente quanto è necessario. Se invece doveste utilizzare un provider di terzi, sareste alla mercé della loro programmazione degli impegni.

Inoltre, se vi occuperete direttamente della manutenzione del vostro server, potrete usufruire di ogni nuova tecnologia non appena viene resa disponibile, senza dovere aspettare che vi venga fornita da un ISP esterno.

Certamente il fatto di possedere un sito Internet implica spese aggiuntive.

Non esiste modo di ovviare a questo aspetto, ma se le vostre intenzioni sono quelle di utilizzare un sito per un lungo periodo di tempo, allora sarà più comodo gestire il sito direttamente. Una volta stabilita la vostra connessione a Internet e avere calcolato i costi di produzione e di manutenzione, raggiungere il vostro pubblico attraverso Internet costa decisamente meno rispetto ai metodi di comunicazione tradizionali, come posta, fax, telefono o visite di persona.

Possibili utilizzi della presenza in Internet

Quello che voi fornirete alla gente attraverso Internet saranno principalmente **informazioni**: informazioni su di voi, sulla vostra azienda o organizzazione, sui vostri prodotti o su qualsiasi altra cosa vogliate far conoscere al pubblico.

In questo senso, il fatto di avere una **vostra presenza in Internet**, vi renderà automaticamente **editori di Internet**.

Potreste anche decidere di offrire alla gente l'opportunità di fare qualcosa di più, che non semplicemente accedere alle informazioni.

Potreste forse desiderare che il vostro sito sia **interattivo**, consentendo lo scambio di informazioni tra voi e i vostri utenti.

O magari riterrete opportuno spingervi ancora più in là e **gestire la vostra attività interamente online**, con una vetrina virtuale.

Ognuno di questi livelli di coinvolgimento nel vostro sito Internet è possibile, ma naturalmente essi richiedono un corrispondente livello di impiego delle risorse.

La **vetrina virtuale** è un tipo complesso di sito Internet e potrebbe essere un'area, nella quale al principiante non è consigliato addentrarsi.

Forse sarà più opportuno iniziare con un sito semplice, pubblicando risorse e lasciando che la vostra presenza in Internet cresca, man mano che sarete disposti a impegnare più tempo e più denaro.

Questa sezione presenta alcuni degli utilizzi più comuni di Internet. Non esiste una procedura di routine. I modi per sfruttare al meglio Internet che qui elenchiamo per

le imprese commerciali, possono essere altrettanto validi anche per altri tipi di organizzazioni.

Potete utilizzare la vostra immaginazione, per decidere quale soluzione sarebbe la migliore per la vostra presenza in Internet.

MOTORE DI RICERCA

Un **Motore di ricerca** è un **software** collegato ad **una pagina web** contenente un **casella** nella quale l'utente introduce i "**termini**" di **ricerca**, composti da **una o più parole**. Questa **successione di termini** viene anche chiamata **query** (in italiano: **interrogazione**).

Questo **strumento di ricerca** è indispensabile, come un **catalogo** in una **biblioteca** o un **indice** in una **enciclopedia**.

Sul **Web**, un indice organizza **milioni di pagine** e **miliardi di parole**. Come un documento normale, la pagina Web può essere dotata di un **indice** seguendo due modalità diverse:

- Utilizzando **parole chiave** che definiscano la **pagina** o il **documento**.
- Utilizzando un **indice** con la **modalità a testo completo**.

Nel primo caso, puoi, per esempio, trovare tutte le pagine che abbiano una relazione con la parola "Turismo" (sarà l'autore della pagina o un utente professionista a inserirla in una determinata categoria).

Nel secondo caso, avrai accesso a tutte le pagine contenenti la parola "turismo".

Possiamo quindi dire che il **motore di ricerca ordina gli oggetti di Internet** (siti, pagine Web, oggetti contenuti in una pagina, ecc.).



Google è attualmente (2007) uno dei **3 motori di ricerca più visitati** negli **USA**, gli altri 2 sono Yahoo e MSN.

La **potenzialità di Internet** aumenta aggiornando periodicamente questi indici. Ogni giorno, nel mondo, vengono creati migliaia di siti, quindi questi indici devono essere **continuamente aggiornati**.

Per eseguire questo, i **motori di ricerca** più grandi **lavorano automaticamente** ed **esaminano il Web** al fine di **scoprire nuovi siti** e **aggiornare le relative pagine**.

Esistono, a disposizione dei "ricercatori": **motori di ricerca, elenchi in linea, funzionalità della guida, selezioni, motori di metaricerca** e **searchbot**.

Come vedremo, esiste una grande quantità di strumenti. Per questa ragione ti sarà utile fare la scelta giusta, imparando il migliore utilizzo dei suddetti strumenti.

Tra questa vasta scelta, il **motore di ricerca**, soprattutto per chi non ha esperienza nel settore, costituisce un **potere inimmaginabile**.

Gli utenti esperti invece potranno rifinire la propria tecnica e imparare a selezionare gli strumenti adatti, ottimizzare le ricerche e interpretarne i risultati.

Ricerca in Internet è semplice.

Puoi lavorare rigorosamente per ottenere la meta che ti sei prefissato senza divagazioni o puoi procedere per mezzo di prove, tentativi ed errori al fine di scoprire l'oggetto misterioso.

Ricercando si impara a ricercare, quindi non esitare a verificare ogni elemento, a sovrapporre richieste, a migliorare i tuoi metodi e il modo in cui ti esprimi.

E' logico che possedere qualche rudimento di cosa siano e come funzionino i motori di ricerca possa agevolare nelle ricerche.

Abbiamo detto che i motori di ricerca aggiornano quotidianamente i loro indici in base ai nuovi siti pubblicati.

Il seguente elenco, presenta i metodi di aggiornamento utilizzati:

Aggiornamento automatico

utilizzando un **programma di ricerca del Web** chiamato "**spider**".

Aggiornamento di un sito eseguito dal relativo autore

che contemporaneamente ordina le proprie pagine in un determinato motore.

Aggiornamento manuale eseguito da utenti professionisti e cyber-giornalisti

il cui lavoro è cercare nuovi siti e aggiornare i database.

Ciascun metodo presenta i propri vantaggi. Sicuramente l'applicazione **helper** può essere di grande aiuto.

Motori come **Infoseek**, **Excite** o **HotBot** possono vantarsi di avere **milioni di pagine** nel proprio indice!

Ma potere non sempre è sinonimo di qualità, poiché più numerose sono le pagine inserite nell'indice, più specifica deve essere la richiesta eseguita dall'utente, altrimenti si viene sommersi da un'infinità di elenchi di risultati.

Fortunatamente, i motori di ricerca possono anche valutare i risultati forniti e, in questo caso, ottenere elenchi in ordine di importanza. Ma il problema sta nel valutare appunto questa importanza.

A volte puoi preferire la selezione manuale eseguita da professionisti.

Inoltre, puoi attribuire grande valore alle guide fornite dai cyber giornalisti che esprimono la propria opinione sui siti visitati e in base a questo prepararne le classificazioni, In questo modo i metodi automatici e umani risultano tra loro complementari.

In cosa consiste un motore di ricerca

Che tu sia un principiante o un professionista, i motori di ricerca costituiscono il **tuo punto privilegiato di entrata nel Web**.

Esaminare da vicino come operano i motori di ricerca, è il modo migliore per comprendere come essi funzionino e chiarire così l'argomento.

Escludendo la distinzione tra i motori di ricerca specifici dell'Italia e più in generale i motori di ricerca che conducono ai siti Web di tutto il mondo, possiamo raggruppare i motori di ricerca nella seguenti categorie:

I motori che inseriscono le pagine Web su indice, per esempio **Excite**, **Altavista** o **HotBot** I motori che organizzano e classificano in titoli e sottotitoli e costituiscono enormi elenchi in linea (per esempio **Yahoo!**, che ricopre sia le versioni italiane nello specifico che quelle a livello mondiale) I motori che valutano e commentano i siti

visitati, per esempio **Virgilio**. Come abbiamo già specificato, esistono numerose tecniche per costituire motori di ricerca o elenchi in linea:

Posizione automatica o indicizzazione.

- Utilizzo di registrazioni manuali create da progettisti ed editor di siti Web.
- Registrazioni e possibilmente valutazioni dei siti da parte di squadre di utenti indipendenti.

Indicizzare pagine Web e siti

Esistono numerosi strumenti che combinano queste diverse possibilità.

I motori di ricerca che utilizzano l'indicizzazione, possono elencare milioni di pagine Web ogni giorno. Ciononostante, il numero di pagine indicizzate non rappresenta l'indicatore essenziale della qualità di un motore di ricerca. Inoltre, i metodi di indicizzazione variano notevolmente da un motore a un altro.

I criteri che permettono la differenziazione delle tecniche utilizzate sono:

- Regolarità di aggiornamento.
- Durata del processo di raccolta pagine per l'indicizzazione.
- Natura dei dati indicizzati: testo, immagini, e altro.

Componenti di un motore di ricerca

Un motore di ricerca è costituito da numerosi componenti:

Lo "**spider**" o "**crawler**" è incaricato di connettere il motore con il server Web, visitare siti, leggere pagine e seguire i collegamenti su queste pagine all'interno dello stesso sito. Una vasta attività! Tutto ciò che lo "spider" trova, verrà utilizzato per compilare l'indice.

Le tecniche più avanzate vengono utilizzate per progettare gli indici dei siti e delle pagine Web. L'indice viene aggiornato regolarmente aggiungendo, eliminando o modificando le relative voci.

Ma esiste uno **squilibrio di tempo** tra ciò che lo spider è riuscito a trovare e quanto di fatto è stato registrato nell'indice.

Il **motore di ricerca, di per sé, è il terzo componente!** Il motore è il componente che analizza una richiesta **facendo riferimento all'indice e in base a questo prepara i risultati**. Inoltre è incaricato di presentare i **risultati in ordine di importanza**.

Esplorare e ricercare

Tenendo presente la difficoltà del tipo di attività (inserire i siti nell'indice e fornire risultati di qualità a milioni di consumatori), **nessun motore può affermare di essere esauriente, offrire il meglio o essere infallibile nei risultati trasmessi**.

Ciò che viene richiesto è **velocità e attendibilità delle informazioni**. È facile capire che l'operazione di un motore di ricerca e le tecniche sono una questione di compromesso. La prestazione varia da sistema a sistema e dipende dalla natura della ricerca da effettuare.

Ciascun sistema possiede punti forti e punti deboli, infatti i motori di ricerca devono essere usati per **sovrapporre e rifinire** le ricerche.

In generale, ciascun motore di ricerca offre:

- Ricerche semplici per mezzo di parole chiave.
- Ricerche complesse utilizzando gli operatori booleani (**AND/ OR /NOT**) e comandi specifici.
- Ricerche per argomento e categorie di siti.

A seconda della prestazione dell'applicazione helper utilizzata, la ricerca può risultare più semplice o difficile. Occorre che tu sia a conoscenza della forza e della debolezza di ciascun motore di ricerca, che a volte dipendono dalla natura della ricerca intrapresa.

Valutare i risultati

La prestazione di un motore di ricerca viene valutata dalla capacità di interpretare la richiesta, valutare i documenti trovati e classificare i risultati della ricerca (bit).

Questo di solito prevede l'assegnazione di una percentuale indicante la corrispondenza della voce con il termine della ricerca. I **primi siti proposti** vengono considerati la **risposta migliore** alla richiesta fatta.

Sempre che, come già sottolineato in precedenza, ci sia la **giusta interpretazione della richiesta** da parte del motore di ricerca (in base al **termine** inserito per la ricerca).

Noterai la velocità di lavoro del motore anche se spesso i risultati includono siti o pagine di poco interesse ai fini della ricerca. Dipende da te ottenere il risultato giusto o sbagliato!

Occorrerebbe comunque essere a conoscenza del fatto che non è compito dei motori di ricerca porre domande aggiuntive per eseguire una ricerca più specifica. È l'utente che deve prendere l'iniziativa e, se necessario, cambiare la richiesta rendendola più adatta e raggiungere così il risultato desiderato. Inoltre i motori non arricchiscono le proprie fonti per mezzo delle tue domande, sebbene queste potrebbero risultare utili per migliorare, in termini di tempo, il servizio.

Sappiamo solo che i **motori di ricerca conservano le richieste eseguite per fornire statistiche interessanti per i consumatori** e che interpretano la natura delle richieste per offrire banner pubblicitari in relazione alle richieste stesse.

In breve, i motori di ricerca sono più intelligenti di quanto si possa pensare.

Puoi contare sulla creatività dei programmatori per poter avere costantemente un miglioramento di servizio in queste applicazioni helper. Possiamo considerare i motori di ricerca come lavoratori instancabili che non si arrendono mai. Digitando semplicemente la parola "Turismo", riceverai migliaia di pagine di risultati.

Se facessi la stessa richiesta a un bibliotecario, sicuramente rimarresti meno soddisfatto e ti verrebbe chiesto di essere più specifico!

Come viene stabilito il **grado di importanza** dei risultati inviati dai motori di ricerca?

Prima di tutto ogni motore adotta il **metodo posizione/frequenza**. Nell'italiano

normale, si usano regole che interpretano la posizione delle parole trovate e la loro frequenza.

La parola "Turismo" avrà maggior peso se contenuta nell'intestazione di una pagina piuttosto che all'interno della pagina Web.

Si osserverebbe, per determinare il grado di importanza, la **frequenza** con cui la parola appare nella pagina stessa qualora la parola fosse, in entrambe le pagine, inserita nel titolo. Anche la **posizione della parola** all'interno della pagina è importante.

Pertanto, il motore di ricerca sa che, se la parola in questione apparirà nella parte alta della pagina, nel titolo o nel primo paragrafo, questa avrà maggior peso rispetto a uno stesso vocabolo che però emerge in fondo alla pagina.

La maggior parte dei motori utilizza il **metodo base** con **una propria variante**.

Questa è una delle ragioni per cui **diversi motori producono risultati altrettanto differenti sulla base della stessa richiesta**.

Come abbiamo già specificato, i **motori di ricerca di Internet** non sono uguali:

- Alcuni registrano più pagine di altri.
- Alcuni aggiornano il proprio indice più velocemente.
- I propri motori hanno un'efficienza diversa.

I motori di ricerca utilizzano altri metodi di differenziazione.

WebCrawler, per esempio, verifica la **popolarità** dei **collegamenti**. Questo significa che darà preferenza a una pagina che presenta numerosi collegamenti, rispetto, ovviamente, a una pagina che ne ha meno.

Qualcuno potrebbe immaginare che i motori di ricerca, associati a directory di siti qualificate, saranno considerati migliori rispetto ai siti valutati manualmente.

I siti visitati da una squadra di professionisti riceveranno valutazioni migliori rispetto ai siti sconosciuti ai professionisti del settore.

DOMINIO

Potremmo incominciare con il definire un dominio internet (in inglese: domain name) come l'equivalente in Internet di un numero di telefono o di un indirizzo, anche se in realtà esso, come vedremo, rappresenta qualcosa di più grave.

I computer comunicano tra loro attraverso l'Internet Protocol (IP) che assegna ad ogni terminale sulla rete un indirizzo composto di quattro numeri separati da punti.

Per esempio, l'indirizzo di google è: 74.125.43.103

I numeri, letti da destra verso sinistra, consentono di individuare: la regione in cui si trova il computer, il provider di servizi di cui si serve, il gruppo specifico di computer a cui la postazione appartiene ed infine l'ultimo numero, che individua la singola macchina nel gruppo.

L'uso degli indirizzi IP, con la diffusione di Internet a livello mondiale e la proliferazione dei siti web, è risultato poco pratico data la difficoltà per l'utente di memorizzare e di operare con queste complesse sequenze di numeri.

Questo problema è stato risolto attraverso il Domain Name System (DNS), che ha consentito di convertire le cifre degli indirizzi IP in parole e denominazioni più facili da ricordare e da utilizzare.

Così, grazie ad un apposito software chiamato browser, il computer è in grado di convertire il **nome di dominio** immesso dall'utente nella stringa di quattro numeri che individua il sito richiesto.

Per esempio nel nostro caso il nome di dominio è: **corsograficoipc.altervista.org**, dove la parola **corsograficoipc.altervista** è il **dominio** vero e proprio (il quale corrisponde al **SLD** o Second Level Domain di cui si parla più avanti), **altervista** è il nostro provider e **.org** è il **TLD** o Top Level Domain.

Una volta digitato il nome, il programma cerca l'indirizzo sui name server, vale a dire i computer su cui sono conservati tutti gli indirizzi IP dei domini registrati.

Se la ricerca nel database del name server ha esito positivo, il computer dell'utente viene collegato all'host richiesto per accedere alle informazioni contenute nel sito web.

Ma vediamo più in particolare la struttura del nome a dominio. In realtà la stringa (sequenza di caratteri) che va digitata nel vostro browser comprende più elementi:

<http://corsograficoipc.altervista.org>

1) Le lettere **http** significano **Hyper Text Transfer Protocol** che è il **linguaggio** usato dal **server** per **comunicare** con i client. Solitamente l'indicazione di questo protocollo appare in automatico sul browser.

Precisiamo: il **server** non è altro che il computer situato presso il nostro provider di presenza (quello per intenderci che sta ospitando il nostro sito cioè l'insieme delle nostre pagine web) che si trova a Firenze (non presso di noi, quindi, che siamo a Imola), mentre i client sono tutti i pc dei visitatori, compreso il tuo che in questo momento ha appena ottenuto dal server i codici HTML che consentono al tuo browser (per esempio Internet Explorer) di leggere questa pagina.

2) spesso vi è l'acronimo **www** (**World Wide Web**) che si riferisce alla nota **piattaforma ipertestuale** comune a tutti i **nomi a dominio**.

3) La terza parola (**altervista**) è il **Second Level Domain (SLD)** e costituisce il **cuore del nome a dominio**. Esso assume in ambito commerciale un'enorme rilevanza distintiva circa l'impresa ed i suoi prodotti o servizi.

4) Vi è poi il **Top Level Domain (TLD)** che nel nostro caso è **'.org'** che è il **suffisso finale** con il quale termina ogni domain name. Esso può essere ricondotto a **due categorie**, una tematica, che esprime i **"generic Top-Level Domains" (gTLD)**, ed una **geografica**, che esprime i **"country code Top-Level Domain" (ccTLD)**.

E' prassi comune non separare mai gli ultimi 2 elementi, in modo da comunicare in modo inequivocabile il TLD di appartenenza. Per esempio se con l'espressione **corsograficoipc.altervista.org** si intende precisamente indicare il **nostro nome di dominio**, ciò significa nel mondo il nostro nome è unico: non può esistere contemporaneamente in nessun altro posto del globo un **corsograficoipc.altervista.org** mentre può esistere un **corsograficoipc.altervista.fr**, un **corsograficoipc.altervista.uk** e così via.

Il titolare di un dominio di 2 livello può creare infiniti sotto-domini di 3 livello semplicemente specificandone il nome alla sinistra del nome di dominio, separato da un punto.

Ad esempio noi al momento non ne abbiamo la necessità, ma potremmo creare **corsograficoipc.altervista.org** e così via. Solitamente ai domini di 2 livello corrisponde una cartella principale detta root entro lo spazio concesso dal provider su una delle sue macchine se l'hosting è frazionato, oppure ad un intero hard disk o ad un intero computer se l'hosting è dedicato.

I domini di 3 livello invece corrispondono quasi sempre a sotto-directories della root principale, all'interno delle quali verranno salvate le pagine web.

Pertanto:

- i **gTLD** sono i domini che dovrebbero distinguere il settore di operatività del sito. Il condizionale è d'obbligo poiché sinora tale destinazione è stata spesso disattesa.

In ogni caso essi sono: .com per i siti commerciali, net per quelli concernenti la rete, . org per quelli di organizzazioni no profit, .gov per quelli delle pubbliche amministrazioni USA, . mil per quelli di enti militari (queste ultime due destinazioni tematiche risultano invece sostanzialmente rispettate) ecc.

- I **ccTLD** sono i domini con estensione finale di due lettere che individuano lo Stato in cui è localizzato l'host server. Qui l'indicazione geografica è solitamente rispettata. Avremo così il suffisso . it per i siti italiani, .de per quelli tedeschi, .es per quelli spagnoli, ecc.

A livello operativo i ccTLD e i gTLD non sono differenti, gli standard tecnici sono gli stessi. Ciò che li distingue è la diversa destinazione e, quindi, il soggetto che ne gestisce la registrazione e l'uso.

Oggi, il governo dei TLD è affidato all'ICANN (Primario Ente Americano per l'assegnazione di nomi di dominio), che ha il ruolo essenziale di coordinare i diversi soggetti che registrano i domini.

I gTLD sono amministrati dai cosiddetti Registrars, società, appunto, di registrazione accreditate dall'ICANN, che operano in regime di concorrenza. Dette società sono autorizzate ad operare grazie a contratti quinquennali stipulati con l'ente (Registrar Accreditation Agreements), che esige specifici parametri tecnico-finanziari per accreditare il soggetto richiedente alla registrazione di TLD.

La Network Solutions Inc., un tempo unico Registrar, conserva il ruolo fondamentale di registry, ossia di gestore materiale del database degli indirizzi e dei nomi di dominio dei gTLD .com, . net e org.

i ccTLDs non sono gestiti direttamente da un organo centrale, ma sono affidati al paese che ne richieda l'attivazione all'ICANN. Ogni Stato decide autonomamente a chi affidare e come gestire il proprio TLD, per così dire, nazionale.

L'ICANN richiede solo che si rispettino dei parametri tecnici specifici e che il servizio sia reso al pubblico con obiettività e senza distinzioni.

Vediamo cosa avviene in Italia per il **dominio geografico .it**.

L'esperienza italiana della diffusione di Internet può essere considerata in linea con l'evoluzione di questo strumento nel mondo, sia per i problemi che sono stati affrontati, sia per le soluzioni adottate.

Anche in Italia si è assistito ad un vero e proprio boom della rete, soprattutto negli ultimi anni novanta.

Usato in principio solo per scopi scientifici dai ricercatori e da pochi altri addetti ai lavori, Internet si è diffuso in maniera vertiginosa e i domain names registrati nel ccTLD .it sono passati dai 153 del 1994 (anno di attivazione dello stesso TLD) agli oltre 500.000 del marzo del 2001.

Le problematiche emerse nell'ambito del ccTLD .it, sia pure in una dimensione "locale", sono le stesse che hanno riguardato i generic Top-Level Domains (.com, .org, ecc.). L'analisi dell'evoluzione tecnica e giuridica della materia dimostra come le esperienze dei singoli paesi, pur con le dovute differenze culturali e strutturali, abbiano portato a soluzioni essenzialmente simili, evidenziando il carattere "globale" di Internet.

L'organizzazione ed il governo del TLD ".it"

In Italia il sistema dei nomi a dominio è amministrato da due enti, la **Naming Authority** (detta anche **NA**) e la **Registration Authority (RA)** che, pur collaborando strettamente, sono autonomi e svolgono ruoli diversi.

Questi soggetti amministrano il TLD **.it** in accordo con le regole per il coordinamento del sistema stabilite a suo tempo dalla **IANA**, Internet Assigned Number Authority (Ente americano).

In particolare, la **Naming Authority italiana** ha la funzione di stabilire le regole delle registrazioni del **dominio .it** e di controllarne l'applicazione, mentre alla **Registration Authority italiana** spetta il compito di **gestire le registrazioni ed i registri operativi del suddetto TLD**.

Va peraltro notato che la stretta connessione tra i due enti è anche fisica, poiché entrambi hanno sede a **Pisa**, presso **l'Istituto per le Applicazioni Telematiche del CNR**.

Questo assetto è il frutto della pluriennale collaborazione del CNR con gli altri organi di ricerca (tra cui la IANA), che negli anni ottanta hanno creato e sviluppato Internet. Sebbene siano stati avanzati dubbi sulla legittimazione giuridica in capo al CNR ad amministrare il DNS in Italia, la situazione è in linea con il sistema di assegnazione dei domini in tutto il mondo. È anzi da ricordare, che la RA è il solo soggetto legittimato ad assegnare i domini .it ai richiedenti, mentre nel caso dei gTLD, i Registrar autorizzati dall'ICANN sono plurimi ed operano in regime di concorrenza tra loro.

Esaminiamo ora le Regole di Naming (RdN) nella loro più recente versione in vigore dal 14.8.2001.

Anzitutto va evidenziata la **natura non normativa** di tali regole, che, per loro espressa disposizione, non possono essere utilizzate “come testo probante”.

La regola generale riguardo all’assegnazione dei domain names . it è stabilita dall’art. 3 delle Regole di Naming: “i nomi a dominio vengono assegnati in uso dalla RA ai richiedenti, seguendo l’ordine **cronologico delle richieste**, come definito dalle Procedure Tecniche di Registrazione”.

Come per gli altri TLDs, anche in Italia le registrazioni dei domain name sono compiute in base ad un criterio di priorità cronologica (“first come, first served”): in altre parole il nome viene assegnato a chi per primo ne faccia richiesta.

La sola contestazione che può quindi essere mossa alla Registration Authority è quella di non aver rispettato l’ordine cronologico delle richieste.

Non vi è alcuna valutazione dei requisiti soggettivi che fanno capo al registrante; questi potrà quindi richiedere (ed ottenere) l’assegnazione di un dominio uguale o simile al nome o al marchio di un altro soggetto.

L’altro punto fondamentale che si evince da questa disposizione è che il nome a dominio è assegnato in uso. Tale diritto sorge per effetto della registrazione e si estingue con il venir meno di questa.