

INTERNETTARE 2000

il WEB, l'ipertesto globale

prof. Gianfranco Rosati

INTRODUZIONE. Per quasi 30 anni Internet è stata usata dall'esercito e dalle università USA per scambiare velocemente dati, nonché per creare grandi "banche dati" comuni. Ma poiché nell'ultimo decennio i personal computer hanno raggiunto una potenza elaborativa paragonabile a quella dei grandi computer di 20-30 anni fa, decine di milioni di nuovi utilizzatori possono ora accedere con facilità a Internet.

SI PUÒ USARE QUALSIASI COMPUTER ? Sì, purché collegato ad una linea telefonica con un "modem", e purché venga stipulato un contratto di abbonamento con un provider, ossia uno dei tanti soggetti pubblici o privati che forniscono l'accesso ai servizi disponibili su Internet.

CHE DIMENSIONI HA INTERNET ?

Un computer che accede oggi a Internet può collegarsi con tutti gli altri computer che a propria volta sono già collegati. Internet è costituita da una rete (priva di centro o di periferia) di milioni di elaboratori collegati tra loro nelle maniere più disparate, situati ovunque nel globo, al punto che ...nessuno ne conosce le effettive dimensioni, che peraltro aumentano di ora in ora !!

CHE LINGUA SI PARLA SU INTERNET ?

Dato che è possibile collegarsi da qualsiasi paese della Terra per inviare o ricevere informazioni, teoricamente su Internet si può trovare di tutto, in qualsiasi lingua... in realtà, la maggior parte di ciò che possiamo trovare è scritta in inglese, in quanto questa è la lingua più diffusa fra coloro che utilizzano Internet più intensamente.

INTERNET è una risorsa:

*

accessibile a tutti con qualsiasi computer...

*

in continua evoluzione...

*

per comunicare con il mondo intero...

*

per ricercare e recuperare informazioni...

*

molto economica per ciò che può fornire.

INTERNET NON È :

*

un programma...

*

una rete unica...

*

uno strumento di proprietà di una società o di un Paese...

*

una struttura per esperti di informatica.

I SERVIZI DI INTERNET:

*

“navigazione” nel World Wide Web (WWW)

*

scambio di posta elettronica (E-mail);

*

partecipazione a gruppi di discussione (newsgroup);

*

trasferimento di archivi e di programmi;

Queste e altre funzioni si integrano fra loro dando origine ad un sistema articolato, completo ed estremamente flessibile.

WWW... “Vuvvuvvù !!?”

il World Wide Web (WWW) è un sistema ipermediale che permette di accedere a miliardi di documenti presenti in Internet, e di spostarsi (“navigare”) fra i link che collegano un documento ad altri....

“sfogliare” o “navigare” Internet significa passare da concetto a concetto, da idea a idea, in un viaggio di scoperta quasi illimitato.

WWW... come funziona ?

La tecnologia WWW, ideata al CERN di Ginevra nel 1990 e disponibile al pubblico via Internet dal 1993, ha permesso di fare un enorme salto di qualità nell’interfaccia utente della rete.

Il WWW si utilizza attraverso un software chiamato “browser” che permette di “sfogliare” con estrema semplicità e immediatezza le pagine ipermediali di Internet.

LA “NAVIGAZIONE” NEL WEB:

il Web consiste di miliardi di pagine... ..organizzate in centinaia di milioni di siti. Un SITO Web è una raccolta di pagine basata sulla logica dell’ipertesto. Ogni PAGINA può avere contenuto multimediale e spesso interattivo. Ciascuno può realizzare e pubblicare un sito Web.

COSA CERCARE NEL WEB:

Chi si collega per la prima volta a Internet e inizia a 'navigare' nella rete ha spesso una idea molto vaga di quanta, e quanto variegata, sia l'informazione raggiungibile. In genere Internet viene considerato uno strumento per accedere a banche dati, cioè a 'depositi' di informazione ragionevolmente ordinata e organizzata: ad esempio il catalogo di una biblioteca, o l'archivio di articoli di un giornale, o una raccolta di dati di borsa. ...ma accanto all'informazione 'organizzata' di una banca dati, Internet offre una ricchezza enorme - anche se di più difficile fruizione - di informazione occasionale, non organizzata, dispersa.

La vera RIVOLUZIONE di Internet consiste nel fatto che CHIUNQUE può mettere informazioni in rete: molti lo fanno già, moltissimi lo faranno nel prossimo futuro.

COSA CERCARE NEL WEB: esempi...

A) un docente può inserire in rete le dispense dei propri corsi, esercitazioni per gli studenti, versioni preliminari dei propri lavori, o magari un quesito sul quale chiedere suggerimenti ai colleghi.

B) un giornalista 'free lance' può utilizzare Internet per farsi conoscere e per diffondere i propri articoli, oltretutto - a sua volta - per ricercare informazioni e contatti.

C) una radio, una televisione, un giornale, una casa editrice, possono inserire in rete programmi, anticipazioni e estratti dei servizi realizzati, oppure aggiornamenti, integrazioni, reazioni alle trasmissioni o ai testi pubblicati.

D) uno studente appassionato di musica, o di letteratura, o di cinema, può diffondere le proprie personali recensioni, o realizzare pagine di informazione sull'autore preferito.

(...e sarebbe sbagliato sottovalutare il potenziale interesse di queste risorse: spesso le notizie raccolte da appassionati sono più complete, più illuminanti e più originali di quanto non venga pubblicato attraverso i canali tradizionali e quelli ufficiali).

QUESTA PLURALITÀ DI OFFERTA INFORMATIVA È L'ESSENZA STESSA DELLA RETE.

COSA È UN BROWSER ?

È un software che consente di...

*

esplorare i documenti presenti sul Web

*

visualizzare il contenuto dei siti

*

navigare tra un sito e l'altro

*

può anche utilizzare programmi associati per riprodurre file audio e video.

I browser più diffusi sono :

*

MICROSOFT EXPLORER

*

NETSCAPE NAVIGATOR

COSA È UN PROVIDER (I.S.P.) ?

Un I.S.P. (Internet Service Provider) è un fornitore di accesso ad Internet.

*

può essere pubblico o privato

*

abbonandovi, potrete connettervi ai

*

potenti elaboratori del vostro ISP

*

collegati permanentemente alla rete

Il provider dunque, da un punto di vista strettamente tecnico, è un insieme di computer connessi da un lato con dei modem pronti a ricevere le chiamate degli utenti, e dall'altro con Internet.

COME SCEGLIERE UN PROVIDER ?

Sono molti i fattori che determinano la **QUALITÀ** di un provider:

1) il numero e il tipo di modem messi a disposizione dei propri utenti;

- 2) la portata e la differenziazione delle linee che lo connettono a Internet ;
- 3) il mutevole rapporto connettività -utenti ;

...inoltre, dato che il costo della chiamata dipende dalla distanza fra voi e il vostro ISP, e dato che ogni ISP offre molti nodi di accesso sparsi sul territorio...

- 4) è opportuno scegliere il provider che abbia diversi nodi nelle principali città italiane e, possibilmente, straniere, per connetterci sempre al costo di una telefonata urbana.

COME RAGGIUNGERE UNA PAGINA WEB ?

...IN TRE MODI:

- direttamente digitando lo specifico indirizzo (URL) che identifica univocamente quella pagina sulla Rete.
- fortuitamente “navigando” attraverso i collegamenti ipertestuali contenuti altre pagine;
- attraverso i motori di ricerca (“search engines”), spesso mediante una serie di approssimazioni successive.

COSA È UNA U.R.L. ?

U.R.L. = Uniform Resource Locator (Localizzatore Universale di Risorse)

è l'indirizzo di un file sul Web. Es:

<http://www.istruzione.it/concorsi/news.htm>

è costituito in genere da quattro parti:

1. protocollo: <http://>
2. server o dominio: www.istruzione.it/
3. percorso: [concorsi/](http://www.istruzione.it/concorsi/)
4. nome del file: [news.htm](http://www.istruzione.it/concorsi/news.htm)

talvolta il percorso o il nome del file non è specificato (es: <http://www.unifi.it>)

COSA È UN MOTORE DI RICERCA ?

Molti sono affascinati da Internet, ma non sanno o cosa cercare o come trovare ciò che cercano.

Esistono centinaia di milioni di pagine WEB e decine di migliaia di gruppi di discussione, ma non esiste un catalogo generale centralizzato.

Una delle prime cose da fare è quindi imparare ad utilizzare uno o più motori di ricerca.

I motori di ricerca (“Search engines”) costituiscono forse l'applicazione Web più conosciuta. Tecnicamente, sono applicazioni utilizzate per individuare documenti sul Web, alle quali è possibile accedere tramite un browser.

Il paradosso è che i motori di ricerca di siti Web ...sono anch'stessi siti Web!! (a proposito, il loro utilizzo è gratuito: i motori di ricerca vivono di pubblicità).

I motori di ricerca Web più noti sono: Excite, Altavista, WebCrawler, Infoseek, Yahoo!, Lycos, ecc... a questi se ne aggiungono sempre di nuovi.

Esistono motori di ricerca "made in Italy" come Virgilio e Arianna, nonché le versioni italiane di alcuni dei maggiori motori americani. Vi sono poi siti sui motori di ricerca stessi, ad es.:

<http://www.motoridiricerca.it/index.htm>

ALCUNI MOTORI DI RICERCA:

Yahoo ! : se non sai il nome,... usa le categorie! Anche in italiano

Northern Light: forse il più potente

Virgilio : il motore di casa nostra

Google : attualmente, il mio preferito...

Altavista : corregge i tuoi errori di ricerca, e traduce i testi dei siti trovati. Anche in italiano

COME FUNZIONA UN MOTORE DI RICERCA ?

Come fa il motore di ricerca ad avere immagazzinato tutte quelle informazioni? e come fa a tenerle aggiornate ?

Naturalmente dietro a tutto ciò non ci sono schiere di impiegati che catalogano e inseriscono manualmente questi dati... il tutto avviene in modo automatico (ma anche su segnalazione di chi vuole promuovere un certo sito): il motore visita le pagine di un sito e le analizza per poter estrarre le informazioni ritenute rilevanti.

Ciascun motore di ricerca utilizza i propri criteri per la catalogazione dei siti e l'estrazione delle chiavi di ricerca.

Tuttavia ALCUNI CRITERI risultano talmente comuni da poter essere considerati standard :

*

il titolo dato a ciascuna pagina Web

*

le parole presenti nel testo, assumendo che le informazioni più importanti vengano inserite all'inizio della pagina.

I motori di ricerca permettono di ricercare parole o combinazioni di parole. Ad es., per cercare le pagine che si occupano di Lewis Carroll, digitiamo le parole 'Lewis' e 'Carroll'.

In molti casi è possibile combinare le parole fornite utilizzando i cosiddetti operatori booleani:

'Carroll' AND 'Alice'

'Carroll' OR 'Alice'

'Lewis Carroll' NOT 'Alice'

La sintassi corretta per utilizzare gli operatori booleani come AND, OR, NOT varia da motore a motore di ricerca.

Secondo uno studio condotto dal centro ricerche della NEC alla fine del 1999:

*

nessun motore di ricerca riesce a indicizzare oltre il 16% delle pagine del Web;

*

l'83% delle pagine ha contenuto commerciale

*

il 6% ha carattere scientifico-educativo

*

solo l'1,5% ha contenuti pornografici

(fonte: <http://wwwmetrics.com>)

I NUOVI MOTORI DI RICERCA:

GLI AGENTI

Gli attuali motori di ricerca presuppongono un utente 'impegnato' nel processo di individuazione della informazione che lo interessa.

Sarebbe utile disporre di un 'segretario' che conosca più o meno i miei interessi, sappia prevedere ed anticipare le mie richieste, compia al posto mio (prendendo autonomamente le decisioni più opportune) le azioni più ripetitive, e magari anche qualcuna di quelle meno ripetitive.

L'idea è quella di disporre di un agente software:

si tratta di incaricare un programma di svolgere determinate ricerche per mio conto, anche a intervalli prefissati, magari chiedendogli di reagire autonomamente ai risultati della ricerca stessa (ad esempio 'filtrandoli' attraverso l'uso di criteri che potrebbero essere difficili o impossibili da impostare direttamente sul motore di ricerca utilizzato).

Se poi il programma 'agente' non è fisicamente legato al mio computer ma è ospitato da un server remoto, potrei affidargli ricerche e compiti da svolgere anche mentre il mio

Il settore degli agenti di ricerca intelligenti sta oggi conoscendo un'evoluzione tale da far sembrare questi primi strumenti solo rozze e primitive approssimazioni ad applicazioni assai più sofisticate, potenti e autonome.

Ecco alcuni degli attuali agenti intelligenti di ricerca fra i siti di commercio elettronico:

<http://www.auctionoctopus.com/>

<http://www.bizrate.com/>

<http://www.bottomdollar.com/>

<http://www.bsilly.com/>

<http://www.mysimon.com/index.anml>

<http://www.pricescan.com/>

<http://www.shopgenie.com/>