

# Internet

La rete delle reti: collega fra loro reti locali, metropolitane, geografiche e singoli computer di tutto il mondo



## Internet

- 1990: 3000 reti e 200.000 calcolatori (detti host)
- 1992: viene collegato il milionesimo host
- Agli esordi il numero di host cresce in modo esponenziale mentre in questi anni si osserva un rallentamento, con un incremento annuo del 6% (sondaggio Nielsen)
- 2002: hanno accesso ad Internet 457 milioni di persone, di cui 174 milioni negli Stati Uniti (sondaggio Nielsen)

## Cosa vuol dire essere in Internet?

- Una macchina è in Internet se utilizza il protocollo **TCP/IP**, diventato ufficiale il 1 gennaio 1983, ha un suo indirizzo IP, ed ha la capacità di spedire pacchetti IP a tutte le altre macchine su Internet
- E' possibile essere in Internet anche in modo temporaneo chiamando un fornitore di servizi Internet (mediante un modem)  
Si parla di Internet Service Provider (ISP): es. Tiscali, Libero, ...

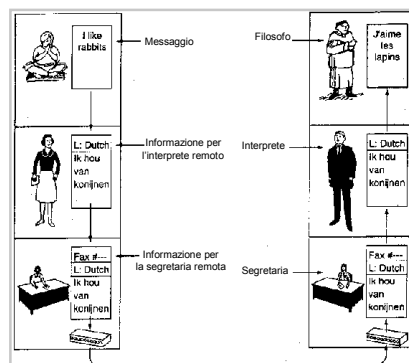
## Cosa vuol dire protocollo?

- La comunicazione tra due calcolatori in una rete è governata da un insieme di regole che prendono il nome di **protocollo**
- Un protocollo fornisce delle funzionalità per
  - indirizzamento (addressing)
  - instradamento (routing)
  - gestione di eventuali errori di trasmissione (error detection, error recovery, sequence control)
  - gestione della velocità di comunicazione (flow control)

## Come si definisce un protocollo?

- Un protocollo "monolitico" che realizzi tutte le funzionalità necessarie per la comunicazione tra elaboratori in rete è difficile da realizzare
- Inoltre, se cambia qualche componente della rete, si deve modificare l'intero protocollo
- Per ridurre la complessità di progettazione la maggior parte dei protocolli è organizzata come una serie di **livelli**  
Il numero dei livelli, il loro nome, le funzionalità differiscono da una rete ad un'altra

## Comunicazione multilivello



### Comunicazione multilivello

Un filosofo desidera comunicare la sua simpatia per l'oryctolagus cuniculus al suo pari. Per fare questo passa un messaggio (in inglese) al suo interprete. L'interprete si è accordato su un linguaggio neutro, l'olandese, e così traduce il messaggio. La scelta del linguaggio è il protocollo di livello 2.

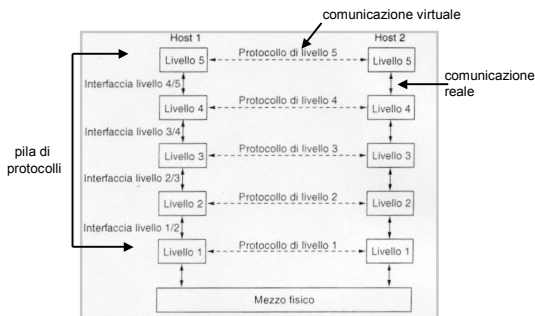
L'interprete consegna il messaggio ad una segretaria per la trasmissione. La segretaria usa un fax (protocollo di livello 1). Quando il messaggio arriva, esso viene tradotto in francese e passato al secondo filosofo.

Gli interpreti possono passare da una lingua ad un'altra purché si mettano d'accordo. Questo non cambia l'interfaccia con i livelli adiacenti. Allo stesso modo le segretarie possono passare dal fax alla posta elettronica senza disturbare gli altri livelli.

### Comunicazione multilivello

- Qualcosa di simile accade nei protocolli di comunicazione tra calcolatori: si ipotizzano dei **livelli** e
  - il livello n di un calcolatore comunica (virtualmente) con il livello n di un altro calcolatore
- In realtà nessun dato viene trasferito da un livello n ad un altro ma passa ad un livello sottostante
  - un protocollo di livello n svolge le sue funzioni usando i servizi forniti dal livello n-1 e fornisce i servizi al livello n+1

### Comunicazione multilivello



### Comunicazione multilivello

- Per ogni coppia di livelli adiacenti esiste una **interfaccia**
- Le convenzioni usate nella conversazione sono il **protocollo**
  - si tratta di un accordo tra i partecipanti su come deve avvenire la comunicazione
- Al di sotto del livello più basso c'è il mezzo fisico che serve per il trasferimento dei dati

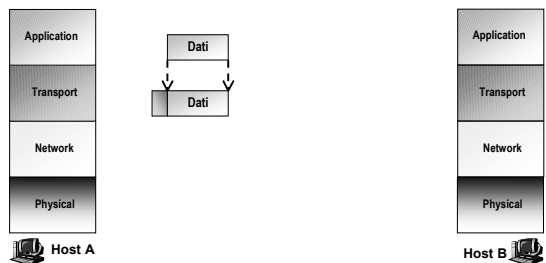
### Comunicazione multilivello: esempio



Il calcolatore A produce un messaggio (Dati) che deve essere inviato al calcolatore B

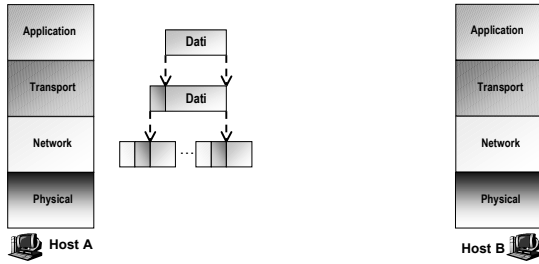
Il messaggio viene prodotto da un programma applicativo nel livello più alto della gerarchia (Application)

### Comunicazione multilivello: esempio



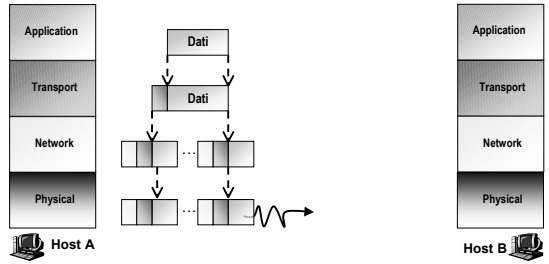
Il livello Application passa il messaggio al livello Transport dove viene inserita un'intestazione (header) che permette di identificarlo

### Comunicazione multilivello: esempio



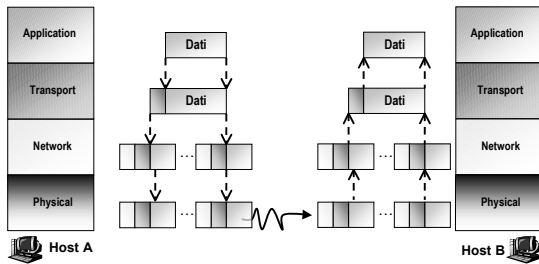
Dal livello Transport il messaggio viene passato al livello Network  
Qui viene suddiviso in parti più piccole (**pacchetti**) e viene aggiunta una intestazione ad ogni pacchetto

### Comunicazione multilivello: esempio



Finalmente si raggiunge il livello più basso (Physical) dove avviene il trasferimento fisico dei dati verso il nodo destinatario B

### Comunicazione multilivello: esempio

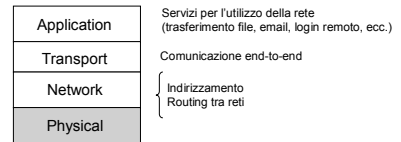


Sul nodo B i pacchetti arrivano al livello più basso e risalgono via via tutti i livelli, con le intestazioni che vengono eliminate di volta in volta

Il messaggio originale viene così ricostruito e ricevuto dal destinatario

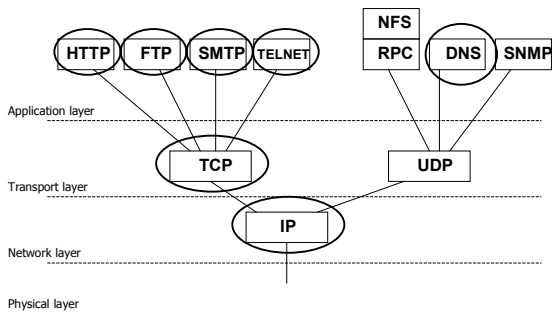
### La famiglia di protocolli TCP/IP

- L'esempio precedente descrive in modo molto astratto un meccanismo di comunicazione del tutto analogo a quello usato per la rete Internet dove si utilizza la famiglia di protocolli TCP/IP
- Si tratta di protocolli organizzati in **livelli concettuali**, ad ogni livello corrispondono determinate funzioni



**NB.** il livello più basso (Physical) è in realtà formato da due livelli distinti, il livello fisico vero e proprio e il livello di accesso alla rete (Data Link) ma non entreremo nei dettagli

### La famiglia di protocolli TCP/IP

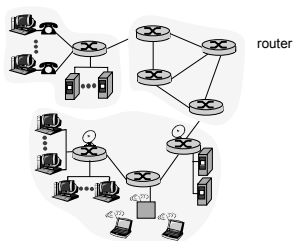


### Network layer: IP

- Internet può essere vista come una collezione di sottoreti diverse (**eterogenee**) connesse insieme (**internetworking**)
- La "colla" che tiene insieme le varie sottoreti è il protocollo **Internet Protocol (IP)**
- Permette di trasportare i dati dalla sorgente alla destinazione, sfruttando la presenza di reti intermedie lungo il percorso

### Network layer: IP

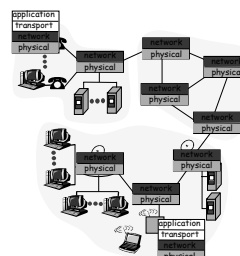
- Una entità di livello Network è presente su tutti i dispositivi collegati in Internet



NB. Una entità è una porzione del sistema operativo che implementa le funzionalità previste da un certo livello

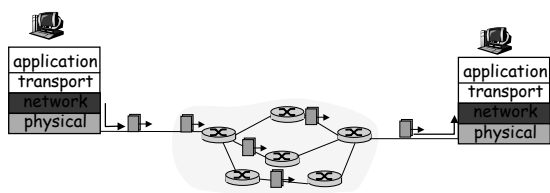
### Network layer: IP

- Una entità di livello Network è presente su tutti i dispositivi collegati in Internet



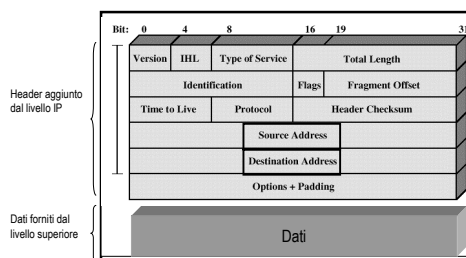
### Network layer: IP

- Trasmissione di tipo packet switching  
I dati, suddivisi in pacchetti, possono seguire percorsi diversi
- I **router** non mantengono informazioni sullo stato delle comunicazioni tra il mittente e il destinatario



### IP: datagram

- Specifica il formato dei pacchetti (**datagram**) spediti sulla rete Internet (lunghezza max 64 KB)
- L'header IP è lungo 20 byte



### IP: datagram

- IP fornisce un servizio di trasferimento di datagram dal calcolatore mittente a quello destinatario
- Il servizio è **connectionless**: ogni datagram è gestito indipendentemente da tutti gli altri
- Il servizio è inaffidabile (best effort): i datagram possono arrivare fuori sequenza oppure possono essere persi

### IP: indirizzi

- Ogni calcolatore collegato ad Internet possiede un indirizzo univoco detto indirizzo IP (**32 bit**)
  - **NetId**: identifica la rete cui il calcolatore è fisicamente collegato
  - **HostId**: identifica il calcolatore

### IP: indirizzi

- I 32 bit di un indirizzo IP sono suddivisi in 4 campi da 8 bit ciascuno

10000000000010100000001000011110

- Di solito si usa una rappresentazione formata da quattro numeri decimali separati da un punto

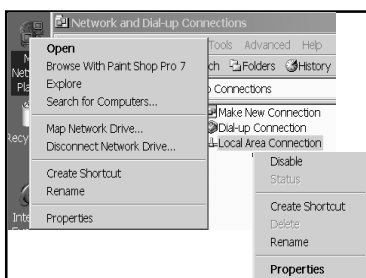
128.10.2.30

### IP: indirizzi

- Gli indirizzi IP devono essere **univoci**
  - per questo motivo è stata istituita una organizzazione, **Internet Assigned Number Authority**, preposta ad assegnare indirizzi IP garantendone l'univocità
- Quando vi collegate ad Internet da casa è il provider che vi assegna un indirizzo IP scegliendolo tra quelli che ha acquistato

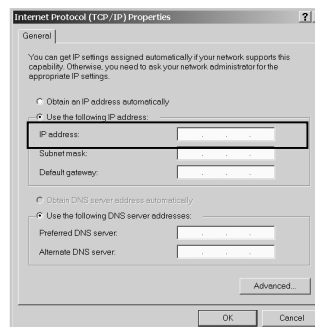
### Configurare l'indirizzo IP in Windows

My Network Places → Properties → Local Area Connection → Properties



### Configurare l'indirizzo IP in Windows

L'indirizzo IP va chiesto all'amministratore della rete

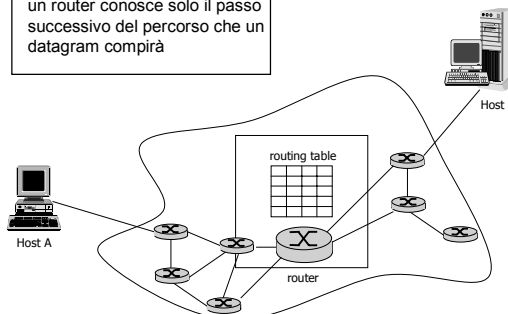


### IP: routing

- IP fornisce anche l'instradamento (**routing**) dei pacchetti tra mittente e destinatario
- Nei router viene mantenuta una tabella di routing che viene usata per trovare il prossimo router o host

### IP: routing

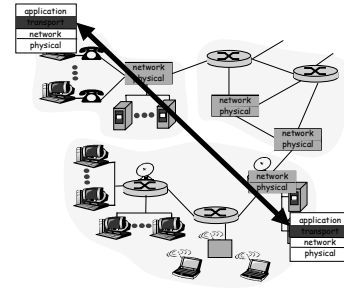
Routing di tipo next-hop:  
un router conosce solo il passo successivo del percorso che un datagram compirà



### Transport layer: TCP

- Il livello Transport è il cuore di tutta la gerarchia di protocolli
- Il suo compito è quello di fornire un trasporto **affidabile** dall'host di origine a quello di destinazione, indipendentemente dalla rete utilizzata
- In Internet il protocollo di questo livello è chiamato **Transmission Control Protocol (TCP)**

### Transport layer: TCP



### Domain Name System - DNS

- Gli indirizzi IP numerici sono difficili da ricordare
- Si usano quindi degli **indirizzi simbolici** che sono più significativi per l'essere umano  
elios.disi.unige.it, samphrey.dcs.ed.ac.uk, developer.netscape.com
- Questi nomi vengono tradotti in indirizzi IP numerici mediante il **Domain Name System**

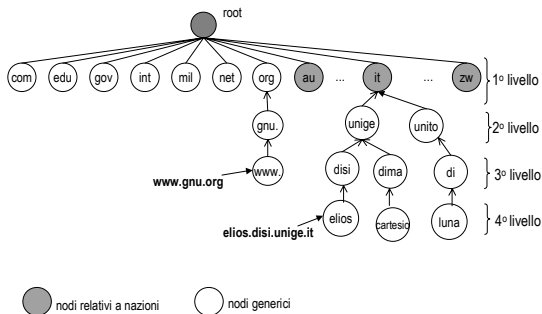
### Domain Name System - DNS

- Gli indirizzi simbolici hanno un formato come quello seguente

...nome5.nome4.nome3.nome2.nome1

### Domain Name System - DNS

- Sono costruiti a partire da uno **schema gerarchico** di nomi basato sul concetto di dominio



### Domain Name System - DNS

- Domini di primo livello (**top level**)

<b>com</b>	aziende
<b>edu</b>	università americane
<b>gov</b>	istituzioni governative
<b>mil</b>	istituzioni militari
<b>net</b>	fornitori d'accesso
<b>org</b>	organizzazioni non-profit
<b>au</b>	Australia
<b>ch</b>	Svizzera
<b>fr</b>	Francia
<b>it</b>	Italia
<b>jp</b>	Giappone
<b>uk</b>	Inghilterra
.....	

### **Domain Name System - DNS**

- Ogni dominio deve essere in grado di "risolvere i nomi" dei calcolatori di sua competenza
- Si usano i **name server** che gestiscono la corrispondenza tra nomi simbolici e indirizzi IP numerici
- Quando un'applicazione deve collegarsi ad una risorsa di cui conosce il nome logico (ad es. albert.unige.it), invia una richiesta al suo name server locale

### **Domain Name System - DNS**

- Il name server locale, se conosce la risposta, la invia direttamente al richiedente. Altrimenti interroga il **name server di top level**. Questi può conoscere l'indirizzo oppure inoltrare l'interrogazione ai suoi figli nella gerarchia
- Si continua con le interrogazioni fino a quando non si ottiene l'indirizzo IP numerico della risorsa
- Quando l'applicazione riceve la risposta crea una connessione TCP con la destinazione, usando l'indirizzo IP appena ricevuto

### **Application layer**

- Si colloca al di sopra del livello Transport ed è il livello nel quale viene svolto il "lavoro utile" per l'utente
- In questo livello si trovano diversi protocolli, alcuni relativi alle applicazioni che usiamo abitualmente in Internet
  - SMTP (Simple Mail Transfer Protocol)
  - FTP (File Transfer Protocol)
  - TELNET
  - HTTP (HyperText Transfer Protocol)

### **Application layer**

- I protocolli del livello Application sono basati sul modello di interazione **client/server**
- Per usare i servizi messi a disposizione mediante questi protocolli bisogna contattare un server
  - Tutte le volte che usate il browser e chiedete delle pagine HTML di un sito, di fatto state contattando un web server remoto
  - Tutte le volte che inviate una e-mail di fatto il mail server del vostro provider contatta il mail server del provider del vostro destinatario

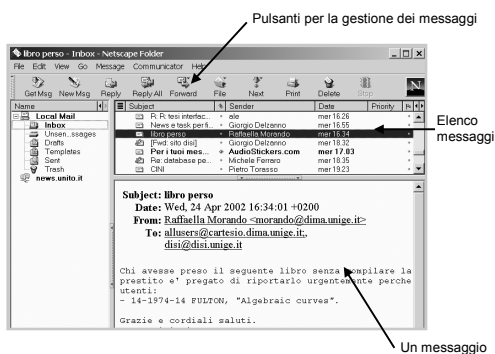
### **Posta elettronica**

- Si basa sul protocollo SMTP (Simple Mail Transfer Protocol) e permette lo scambio dei messaggi tra gli utenti collegati alla rete
- È necessario fornire
  - l'indirizzo del mittente
  - l'indirizzo del destinatario
  - il corpo del messaggio
- Gli indirizzi devono avere un formato ben preciso  
Esempio: ribaudo@disi.unige.it

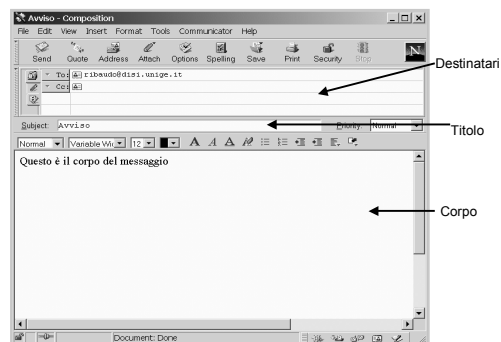
### **Posta elettronica**

- Esistono vari programmi applicativi per la gestione della posta elettronica
  - Microsoft Outlook
  - Messenger di Netscape
- Questi programmi permettono di inviare nuovi messaggi, visualizzare l'elenco dei messaggi ricevuti, salvarli, rispondere

### Posta elettronica: Messenger di Netscape



### Posta elettronica: Messenger di Netscape

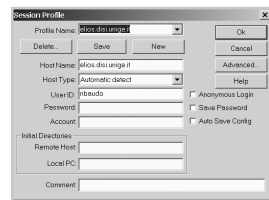


### Trasferimento file

- Si basa sul File Transfer Protocol (FTP)
- Permette di collegarsi a siti remoti per prendere (download) / salvare (upload) file
- L'accesso può essere riservato (tramite login e password) oppure aperto a tutti (si parla di anonymous ftp)

### Trasferimento file

- Esistono vari programmi applicativi che permettono il trasferimento dei file



- Si può attivare il trasferimento file anche usando il browser

### Collegamento remoto

- Telnet
  - permette di collegarsi a calcolatori che si trovano in località remote per lavorare interattivamente su di essi



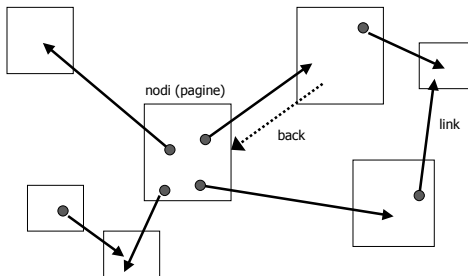
### World Wide Web (WWW)

- Si basa sul protocollo HTTP (HyperText Transfer Protocol) ed è la vera novità degli anni '90
- Sviluppato presso il CERN di Ginevra è il più potente mezzo di diffusione telematica di documenti elettronici
- Mezzo di comunicazione globale, interattivo, multimediale e ipertestuale ha cambiato radicalmente il modo di comunicare e di lavorare



### World Wide Web

- I documenti sono organizzati in forma **ipertestuale**



La lettura avviene secondo legami associativi

### World Wide Web

Il linguaggio HTML che abbiamo visto nelle prime lezioni è il linguaggio che è stato proposto per scrivere gli ipertesti per il web



### Terminologia per gli ipertesti sul web

- **browser**: programma applicativo per navigare in rete
- **page** (pagina): singolo "foglio" di un ipertesto
- **home-page**: "punto di ingresso" di un sito web
- hotspot, **hotword**: porzione di una pagina che, se selezionata, permette di raggiungere un altro punto dell'ipertesto o una nuova risorsa

### Terminologia per gli ipertesti sul web

- **Web**: insieme di tutti gli ipertesti
- **FAQ**: domande ricorrenti su un certo argomento
- **motore di ricerca**: sito che permette di cercare documenti e siti, di solito mediante inserimento di parole chiave
- **portale**: sito Web che offre risorse e servizi; è il punto di ingresso verso altri siti

### Client / Server per il web

- Il WWW usa il modello client/server per lo scambio delle informazioni
- Abbiamo
  - il **client** (browser), programma applicativo che "gira" sull'elaboratore dell'utente
  - il **server**, programma applicativo che "gira" sull'elaboratore del fornitore di informazioni (provider)
  - la **rete**

### Client / Server per il web

- Ogni utente può richiedere delle informazioni attraverso il suo programma client
- La richiesta "viaggia" attraverso la rete fino a raggiungere l'elaboratore server
- Il server interpreta la richiesta e inoltra all'elaboratore client un file contenente le informazioni desiderate (oppure restituisce un messaggio di errore se il file richiesto non esiste sul server)

NB: questa comunicazione avviene usando gli indirizzi IP di client e server!

### Usare il browser

- Per navigare nel web si usano dei programmi applicativi detti browser



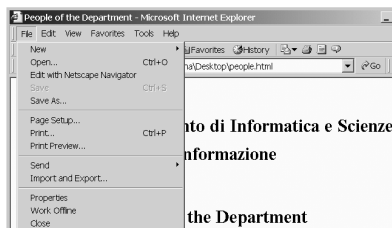
### Usare il browser

- Uno dei menu più importanti del browser è il menu **File** che permette di scegliere quali documenti visitare (Open ...)



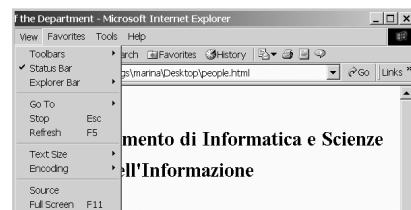
### Usare il browser

- Si possono salvare i documenti HTML usando il comando **Save As**; si possono stampare usando il comando **Print**



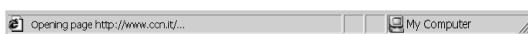
### Usare il browser

- Sotto il menu **View** si trovano i comandi per scegliere quali barre visualizzare e per vedere il codice HTML delle pagine (Source)



### Usare il browser

- Lo stato della connessione si legge nella **barra di stato** in basso
  - Looking up host
  - Contacting host
  - Host contacted waiting for reply
  - Opening page ...



### Usare il browser

- Usando i pulsanti della barra degli strumenti si possono eseguire delle azioni in modo veloce
- Quando si usa il pulsante **Back**, il browser visualizza la pagina prendendola dalla sua memoria **cache** (anziché contattare nuovamente il server)



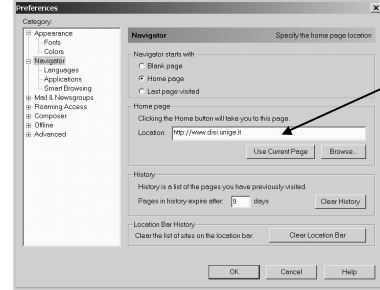
## Usare il browser


- Gli indirizzi per “trovare” i documenti nel Web sono detti **Uniform Resource Locator (URL)** e li identificano in modo univoco
- Di solito sono formati dalla parola **www** seguita dal **nome simbolico** del nodo su cui si trova il server, seguito dal **pathname** del file sul server
- Vengono scritti nella barra degli indirizzi



## Configurare il browser (Netscape Navigator)

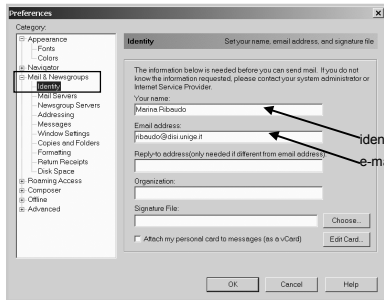
### Edit → Preferences



scelta della pagina iniziale che viene visualizzata quando si preme il pulsante 

## Configurare il browser (Netscape Navigator)

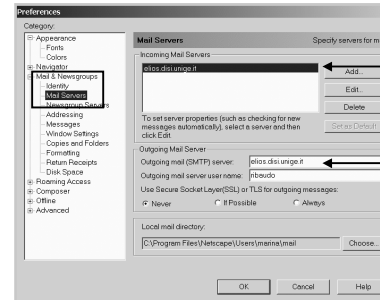
### Edit → Preferences



identità del mittente e-mail del mittente

## Configurare il browser (Netscape Navigator)

### Edit → Preferences

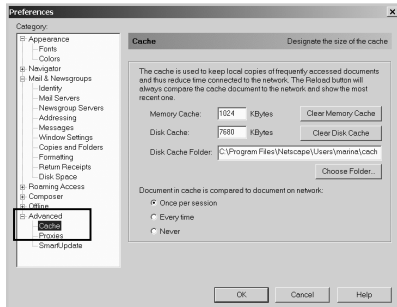


nome logico del mail server in entrata  
nome logico del mail server in uscita

Queste informazioni le fornisce il provider quando si richiede un indirizzo di e-mail

## La memoria cache del browser

### Edit → Preferences



## Motori di ricerca

*A program that searches documents for specified keywords and returns a list of the documents where the keywords were found. Although search engine is really a general class of programs, the term is often used to specifically describe systems like Alta Vista and Excite that enable users to search for documents on the World Wide Web.*

*Typically, a search engine works by sending out a spider to fetch as many documents as possible. Another program, called an indexer, then reads these documents and creates an index based on the words contained in each document. Each search engine uses a proprietary algorithm to create its indices such that, ideally, only meaningful results are returned for each query.*

dal sito [www.webopedia.com](http://www.webopedia.com)

### Motori di ricerca

- Permettono di “districarsi” nel deposito caotico di dati presenti nella rete
- **Indicizzano** i documenti mediante algoritmi sofisticati che analizzano le parole presenti nel documento stesso
- Usano dei **sofbot** (o spider) per esplorare il web alla ricerca dei documenti

### Motori di ricerca

- La ricerca dei documenti avviene quasi sempre mediante **parole chiave**
  - l'utente deve specificare in pochi termini le proprie esigenze informative
- Problema: molto spesso si ottengono troppe risposte (**sovraccarico informativo**)
- Per diminuire il numero di risposte si deve **specializzare** ulteriormente la domanda

### Motori di ricerca

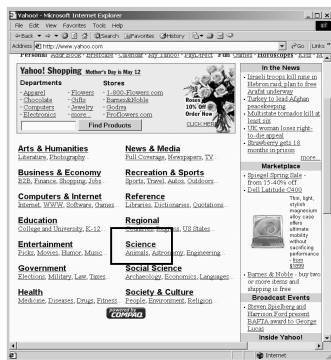
- Si possono usare operatori booleani
  - AND, OR, NOT
- Operatori di vicinanza
  - NEAR
- Espressioni esatte
  - “.....”

**NB:** ogni motore di ricerca adotta una sua sintassi convenzionale. Si deve selezionare la voce **Advanced Research** per vedere le opzioni messe a disposizione

### Motori di ricerca

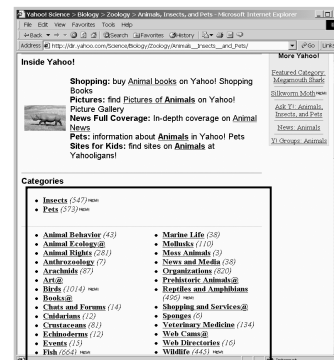
- www.yahoo.com
- www.altavista.com
- www.lycos.com
- www.google.com
- www.excite.com
- www.virgilio.it

### www.yahoo.com

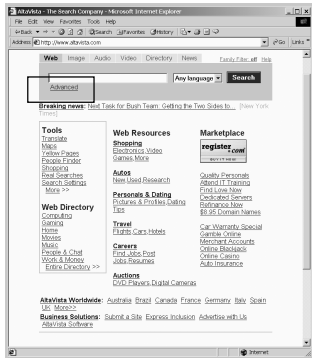


Struttura a directory

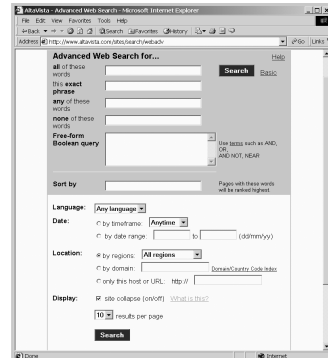
### www.yahoo.com



www.altavista.com



www.altavista.com

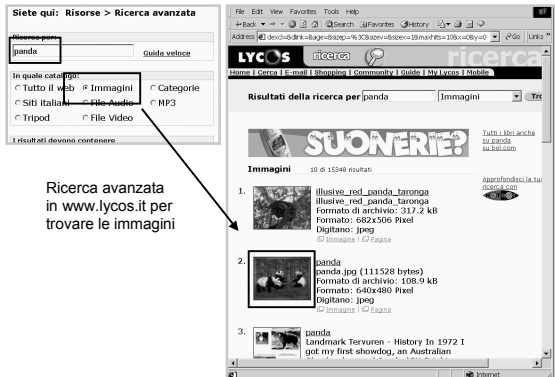


Modulo per la ricerca avanzata

www.lycos.it

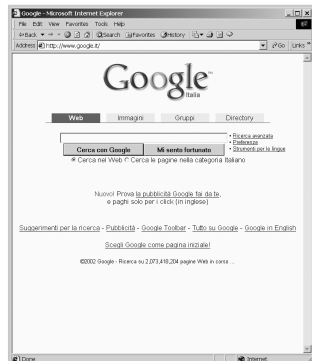


www.lycos.it



Ricerca avanzata in www.lycos.it per trovare le immagini

www.google.it



www.google.it



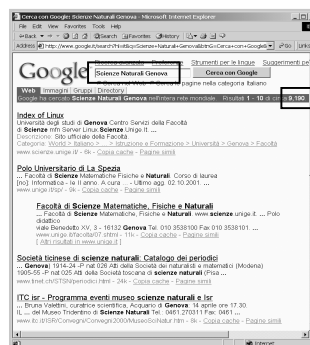
Ricerca per "Naturali" → 430.000 documenti trovati

www.google.it



Ricerca per "Scienze Naturali" → 96.900 documenti trovati

www.google.it



Ricerca per "Scienze Naturali Genova" → 9.190 documenti trovati

### Motori di ricerca

Quando fate una ricerca  
dovete scegliere con cura le  
parole chiave: trovare migliaia  
di documenti equivale a non  
trovarne nessuno



### ... per concludere ...

Internet è molto di più quello  
che abbiamo visto, offre  
opportunità immense ...  
occhio al  
DIGITAL DIVIDE



Per informazioni su alcuni servizi offerti da Internet potete leggere Internet flash di R. Viscardi, Apogeo