

Ministero dell'Istruzione, dell'Università e della  
Ricerca Servizio Automazione Informatica e  
Innovazione Tecnologica

## **Modulo 12**

### Navigare in Internet

**ForTIC**

Piano Nazionale di Formazione degli Insegnanti sulle  
Tecnologie dell'Informazione e della Comunicazione

**Percorso Formativo C**

Materiali didattici a supporto delle attività  
formative  
2002-2004

**Promosso da:**

- Ministero dell'Istruzione, dell'Università e della Ricerca, Servizio Automazione Informatica e Innovazione Tecnologica
- Ministero dell'Istruzione, dell'Università e della Ricerca, Ufficio Scolastico Regionale della Basilicata

**Materiale a cura di:**

- Università degli Studi di Bologna, Dipartimento di Scienze dell'Informazione
- Università degli Studi di Bologna, Dipartimento di Elettronica Informatica e Sistemistica

**Editing:**

- CRIAD - Centro di Ricerche e studi per l'Informatica Applicata alla Didattica

**Progetto grafico:**

- Campagna Pubblicitaria - Comunicazione creativa

In questa sezione verrà data una breve descrizione del modulo.

Gli scopi del modulo consistono nel mettere in grado di:

- Usare particolari tipi di *file* e protocolli e di meccanismi di accesso remoto.
- Usare *tool* e utilità di *Internet*, installare e configurare *browser*, lettori di *newsgroup*, *client* di *chat*.

Il modulo è strutturato nei seguenti argomenti:

- **Usare *Internet***
  - Descrivere le componenti di un URL e il loro significato.
  - Usare *tool* e utilità *Internet* quali: *e-mail*, *browser*, motori di ricerca, *newsgroup*, *list server*, *chat*, trasferimento di *file*.
- **Installazione e configurazione di applicazioni**
  - Installare e configurare un *browser* per *Internet*.
  - Installare e configurare un *browser add-ons and plug-ins*.
  - Installare e configurare un lettore di *newsgroup*.
  - Installare e configurare un *client* di *chat*.

# Introduzione

## Protocolli

Luciano Margara

### Internet: le basi

Una rete telematica è un insieme di nodi e di connessioni.

- I nodi sono unità di elaborazione. Con il termine unità di elaborazione intendiamo una vasta gamma di dispositivi tra i quali troviamo *personal computer*, *computer* portatili, palmari, telefoni cellulari, eccetera.
- Le connessioni di una rete telematica sono mezzi fisici tramite i quali i nodi possono comunicare ovvero scambiarsi informazione. Esempi di connessioni: cavi elettrici e telefonici, campi radio, fibre ottiche.

Le reti telematiche possono essere classificate in base alla loro dimensione in:

- Reti personali (*Personal Area Network* PAN). Reti che collegano un numero ridotto di unità. Esempio: una rete casalinga che collega due *personal computer* e una stampante.
- Reti locali (*Local area network* LAN). Sono reti di dimensioni ridotte. Esempio: rete di *personal computer* di un qualsiasi ufficio.
- Reti metropolitane (*Metropolitan Area Network* MAN). Sono reti progettate per coprire un'area geografica limitata. Esempio: rete di calcolatori del comune di Bologna.
- Reti geografiche (*Wide area network* WAN). Sono reti progettate per collegare siti o reti geograficamente distanti.
- Inter-reti. Sono collezioni di LAN collegate da WAN. Esempio: **Internet**.

**Internet** è quindi in prima approssimazione una rete di reti. Ovvero è una rete che collega altre reti. Le dimensioni di **Internet** sono enormi. Basti pensare che **Internet** è in grado di collegare milioni di calcolatori e di altri dispositivi dislocati sulla superficie del pianeta in più di cento paesi.

Date le dimensioni di **Internet** e la varietà dei dispositivi ad essa collegati il primo problema da risolvere per rendere possibile la comunicazione è quello di definire un linguaggio comune.

*Computer* diversi usano sistemi operativi, codici di caratteri, strutture di dati, che possono essere anche molto diversi. Per permettere loro di comunicare è necessario definire un insieme di regole comuni di comportamento. Queste regole nell'ambito della telematica prendono il nome di protocolli.

## TCP/IP

Un protocollo di comunicazione stabilisce regole comuni che un *computer* deve conoscere e alle quali deve attenersi per inviare e ricevere informazioni attraverso la rete. Un protocollo deve specificare ad esempio in che modo va codificato il segnale, in che modo far viaggiare i dati da un nodo all'altro, in che modo assicurarsi che la

trasmissione sia andata a buon fine.

Nel caso di **Internet**, che interconnette milioni di *computer* e di sottoreti, basati su ambienti operativi e architetture *hardware* diverse, tali protocolli debbono rispondere ad esigenze particolarmente complesse.

L'insieme di protocolli su cui si basa il funzionamento di **Internet** si chiama *TCP/IP* che è un acronimo di *Transmission Control Protocol/Internet Protocol*. *TCP* e *IP* sono due tra i più importanti protocolli di *TCP/IP*, per questa ragione l'intera *suite* di protocolli *TCP/IP* ha questo nome.

*TCP/IP* è la chiave del successo di **Internet**.

*TCP/IP* è indipendente dal modo in cui la rete è fisicamente realizzata: una rete basata su *TCP/IP* può appoggiarsi indifferentemente su una rete locale *Ethernet*, su una linea telefonica, su un cavo in fibra ottica, su una rete di trasmissione satellitare. *TCP/IP* nasconde all'utente la realizzazione fisica della rete che sta usando.

*TCP/IP* è stato progettato per sfruttare al meglio le risorse di comunicazione disponibili; permette un indirizzamento efficiente e sicuro dei *computer* collegati, anche se questi sono diversi milioni; garantisce con la massima sicurezza il buon fine della comunicazione; infine permettere lo sviluppo di risorse e servizi di rete evoluti e facilmente utilizzabili dall'utente.

*TCP/IP* è un *open standard*, le cui specifiche sono liberamente utilizzabili da chiunque. Questo ha permesso il rapido diffondersi di implementazioni per ogni sistema operativo e piattaforma esistente, implementazioni spesso distribuite gratuitamente o integrate in modo nativo nel sistema stesso.

Ciò che viene comunemente indicato come *TCP/IP*, in realtà, è costituito da un vero e proprio insieme di protocolli di comunicazione, ognuno con un compito specifico e organizzati in maniera gerarchica per livelli.

- Livello delle applicazioni. Gestisce i servizi di rete per l'utente e dunque è la fonte e la destinazione finale di ogni transazione di rete.
- Livello di trasporto. Gestisce l'organizzazione dei dati ai fini della trasmissione e ha il compito di controllare che la comunicazione di un certo blocco di dati sia andata a buon fine, e di ritrasmettere quello che eventualmente è andato perso.
- Livello di rete. Gestisce l'indirizzamento dei *computer* e l'instradamento dei dati.
- Livello fisico e di collegamento dati. Gestiscono l'uso dei cavi e l'invio dei segnali fisici sulla rete.

In questo modulo ci concentreremo sull'insieme di protocolli di livello applicazione ovvero su quell'insieme di regole che permettono all'utente di usufruire di servizi telematici quali la posta elettronica e la navigazione del *Web*.

A questo punto possiamo dare una definizione di **Internet** più specifica e dettagliata di quella data in precedenza.

**Internet** è un sistema informativo globale:

- logicamente collegato da un unico spazio di indirizzi, cioè ogni macchina connessa a **Internet** è individuata logicamente da un indirizzo assegnato seguendo le specifiche del protocollo *IP* di *TCP/IP* (Indirizzo *IP*);
- che supporta comunicazioni usando i protocolli *TCP/IP*;
- che fornisce, usa e rende accessibile, in modo pubblico e/o privato, servizi ad alto livello (Esempio: posta elettronica, *chat*, *Web*) basati su *TCP/IP*.

Risulta evidente come *TCP/IP* sia l'elemento cardine per la definizione ed il funzionamento di **Internet**.

Qualunque dispositivo per essere connesso a **Internet** deve aver installato e correttamente configurato il *software* relativo ai (che implementa i) protocolli *TCP/IP*.

Nel seguito prenderemo in esame alcuni dei principali protocolli di livello applicazione nella gerarchia *TCP/IP* ed i servizi a loro associati.

## Servizi di Internet - Telnet, Ftp, Gopher

**Telnet**, ftp e **Gopher** sono tre protocolli di livello applicazione nella gerarchia di protocolli *TCP/IP* (ricordiamo che *TCP/IP* è l'insieme dei protocolli che regolano il funzionamento di **Internet**).

**Telnet** e ftp conservano ancora una grande utilità nonostante il loro utilizzo sia stato ridimensionato nel corso degli ultimi anni. **Gopher** ha quasi totalmente perso la sua utilità dopo l'avvento del *Web*.

- **Ftp** è ancora oggi il metodo più efficace e veloce per trasferire *file* da un *computer* ad un altro.
- **Telnet** è ancora oggi utilizzato per accedere (da una qualunque macchina collegata a *Internet*) a numerosi *computer* che accettano solo connessioni in modalità terminale.
- **Gopher** è stato il primo tentativo di organizzare, attraverso un'unica interfaccia (a caratteri o testuale), l'enorme quantità di informazioni disponibili su **Internet**, poi superato dal *Web*.

## Ftp

**Ftp** (acronimo di *File Transfer Protocol*) è il protocollo che permette di trasferire *file* da un *host* ad un altro attraverso **Internet**.

Memorizzati negli *hard disk* degli *host* connessi a **Internet** c'è una enorme quantità di programmi, immagini digitali, suoni, eccetera, molti dei quali di pubblico dominio.

Il sistema più efficace per trasferire questi *file* da un *computer* ad un altro è usare un'applicazione (*client*) basata sul protocollo FTP.

Le macchine che entrano in gioco in una sessione ftp sono due: la macchina locale (quella sulla nostra scrivania) e la macchina remota (quella sulla quale o dalla quale vogliamo trasferire *file*).

In una sessione ftp è possibile effettuare due operazioni principali, quella di *download* e quella di *upload*.

L'operazione di *download* è l'operazione con la quale si trasferiscono *file* dalla macchina remota alla macchina locale.

L'operazione di *upload* è l'operazione con la quale si trasferiscono *file* dalla macchina locale alla macchina remota.

Sia l'operazione di *download* che quella di *upload* possono essere eseguite solo sulle macchine *remote* che accettano sessioni ftp.

Non tutti i *file* (programmi, immagini, filmati, eccetera) memorizzati su *host* abilitati a sessioni ftp sono prelevabili o prelevabili gratuitamente. Spesso le università, i centri di ricerca, e anche i singoli utenti privati mettono a disposizione *file* di pubblico dominio e trasferibili con diverse modalità di pagamento.

Programmi *shareware*: è richiesto il pagamento di una modica somma di denaro.

Programmi *freeware*: *software* completamente gratuito.

Oltre al *software shareware* e *freeware* esistono altre categorie di *software*. Per esempio *software* di tipo *giftware* (richiede come pagamento un versamento volontario anche non necessariamente in denaro) e di tipo *cardware* (richiede come pagamento la spedizione di una cartolina!)

Per trasferire un *file* usando il protocollo ftp occorre avere a disposizione un'applicazione ftp chiamata *client* ftp (ce ne sono diverse anche di tipo *freeware*) che viene attivata sulla macchina locale e occorre conoscere l'indirizzo *IP* o il nome logico della macchina remota.

Ci sono due modalità di collegamento ad una macchina remota: ftp anonimo, e ftp con *account*.

Il trasferimento di *file* tramite ftp anonimo è quello tradizionalmente utilizzato per il prelievo di *file* ad accesso pubblico presso università, enti, società. Consiste in un *login*, ovvero nell'ingresso in un *computer* remoto, effettuato senza disporre presso di esso di un proprio codice utente e di una propria *password*, quindi anonimamente. In questa modalità non avremo, per ovvi motivi di sicurezza, pieno accesso al *computer* remoto; potremo quindi entrare solo in determinate *directory* - tipicamente nella *directory* chiamata *pub* (ovvero *public*) e nelle sue *sottodirectory* e potremo solo leggere alcuni *file*, ma non cancellarli, spostarli o modificarli.

L'utilizzazione di ftp con *account*, invece, dà pieno accesso ad una determinata *directory* del sistema remoto, nella quale potremo inserire, modificare e cancellare *file*, proprio come se fosse una *directory* del nostro *hard disk*. Di norma è riservata ai dipendenti dell'università, dell'ente o della società che ospita il *server* ftp, oppure ai loro collaboratori, oppure ancora ai loro clienti. Se, ad esempio, decidete di pubblicare su **Internet** una vostra pagina *Web* acquistando lo spazio presso un *Internet provider*, quest'ultimo con ogni probabilità vi concederà un *account* ftp e una *password* personale.

Esistono *client* ftp dotati di interfaccia a caratteri e *client* ftp dotati di interfaccia grafica.

## Programmi con interfaccia a caratteri

I *client* FTP con interfaccia a caratteri possono sembrare di difficile utilizzo, ma sono molto efficienti e versatili.

Vediamo adesso una sessione ftp con *account* eseguita dall'utente margara utilizzando un *client* ftp con interfaccia a caratteri.

Ci colleghiamo ad una macchina del dipartimento di Scienze dell'Informazione dell'Università di Bologna il cui nome logico è alice.cs.unibo.it.

Dopo aver impartito il comando:

```
ftp alice.cs.unibo.it
```

otteniamo la seguente schermata:

```
ftp alice.cs.unibo.it
Connected to alice.cs.unibo.it.
220 alice.cs.unibo.it FTP server (Version 6.2/OpenBSD/Linux-0.10) ready.
Name (alice.cs.unibo.it:margara): margara
331 Password required for margara.
Password:
230- Linux alice 2.2.19pre17 #5 Wed Apr 4 15:05:17 CEST 2001 i686 unknown
230- SCIENZE DELL'INFORMAZIONE - UNIVERSITA' DI BOLOGNA -
CS.UNIBO.IT
230- -----
230- cluster: Linux CS
230- -----
230-(lm 10.01.2003)
230- -----
230 User margara logged in.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp>
```

A questo punto il *computer* alice.cs.unibo.it è pronto per l'operazione ftp vera e propria.

Vediamo adesso la schermata corrispondente al *download* del *file* pippo.txt da alice.cs.unibo.it.

```
ftp> get pippo.txt local: pippo.txt remote: pippo.txt 227 Entering Passive Mode
(130,136,2,13,17,206) 150 Opening BINARY mode data connection for
'pippo.txt' (0 bytes). 226 Transfer complete.
```

L'operazione inversa (*upload*) viene eseguita usando il comando *put* come segue.

```
ftp> put pippo.txt local: pippo.txt remote: pippo.txt 227 Entering Passive Mode
(130,136,2,13,17,100) 150 FILE: pippo.txt 226 Transfer complete. ftp>
```

## Programmi con interfaccia grafica

Le interfacce grafiche hanno definitivamente avuto la meglio nel mondo del *software* di rete, così anche per il prelievo di *file* sono ormai disponibili programmi a base di icone e *mouse*. Ce ne sono diversi, per tutti i sistemi operativi. Due *client* ftp molto usati sono quelli inclusi nei *browser* *Internet Explorer* e *Netscape Navigator*.



Il *client* ftp di *Netscape*, è fortemente integrato con le altre funzionalità del programma. Per collegarsi con un sito ftp tramite *Netscape* è sufficiente inserire l'indirizzo (in gergo tecnico **url**) del sito che vogliamo raggiungere (se la macchina remota è ad esempio *alice.cs.unibo.it*, la relativa url sarà *ftp://alice.cs.unibo.it* dopo aver fatto click con il *mouse* sul comando *Open Page* presente all'interno del menu *File*.

Notare che è possibile inserire la url completa di *path* (ad esempio *ftp://alice.cs.unibo.it/pippo.txt*), così da saltare direttamente alla *subdirectory* che ci interessa.

Per trasferire il *file* sul proprio *hard disk*, a questo punto, non si deve fare altro che un click su *Save File* e attendere che il trasferimento sia completo.

## Protocolli, servizi e applicazioni

A questo punto è necessario spendere due parole per cercare di chiarire il legame che esiste tra i concetti di protocollo, servizio ed applicazione.

I termini protocollo, servizio ed applicazione sono ricorrenti in ambito informatico e rivestono una notevole importanza. Quando si parla di **Internet** il significato di questi termini può essere sinteticamente descritto come segue.

- **Protocollo:** insieme di regole e di convenzioni da rispettare perché due calcolatori possano comunicare.
- **Servizio:** prestazione fornita da una macchina servente (*server*) a una macchina cliente (*client*) che ne fa richiesta.
- **Applicazione:** programma.

Spesso lo stesso termine può, a seconda del contesto in cui è usato, rivestire sia il significato di protocollo che di servizio o di applicazione. Facciamo un esempio: ftp.

ftp è un protocollo di livello applicazione in *TCP/IP*. In particolare è il protocollo che regola il trasferimento di *file* da un *computer* ad un altro.

ftp è il servizio fornito da alcuni *host* di rete chiamati *server* ftp. Collegandosi a un *server* ftp è possibile fare *upload* e *download* di *file*.

**Ftp** è un'applicazione. Ovvero è il programma cliente che permette ad una macchina locale di usufruire del servizio ftp messo a disposizione da un *server* ftp seguendo il protocollo ftp.

## Telnet

**Telnet** è un protocollo di livello applicazione nella gerarchia *TCP/IP* che permette all'utente di visualizzare sulla macchina locale davanti alla quale è seduto un terminale agganciato ad una qualsiasi macchina remota sulla quale (in genere) l'utente possiede un *account*. Utilizzando il terminale così attivato, l'utente può accedere alle risorse della macchina remota come ci fosse seduto davanti. Con telnet, attraverso una interfaccia a caratteri è possibile agire sulla macchina remota proprio come se fosse il nostro *computer* locale, sfruttando tutte le risorse che il sistema mette a disposizione. Ad esempio l'utente può mandare in stampa un *file* sulla stampante collegata alla macchina remota oppure leggere la propria posta elettronica.

L'utilizzazione di telnet è molto semplice. Disponendo di un *client* telnet è sufficiente aprire una finestra di testo e come in una sessione ftp digitare:

**telnet alice.cs.unibo.it**

dove, come al solito, alice.cs.unibo.it è la macchina remota sulla quale vogliamo attivare una sessione telnet. L'indirizzo del *computer* cui ci si vuole collegare può essere un indirizzo numerico (ad esempio: 130.136.1.110) o un nome simbolico (ad esempio: alice.cs.unibo.it).

Dopo pochi secondi appare una schermata di benvenuto, e viene avviata la procedura che consente di farci riconoscere e di accedere alle risorse che ci interessano (procedura di *login*).

Molto spesso per portare a termine una connessione telnet è necessario avere il permesso di operare sul *computer* remoto. Appena stabilita la connessione, il *computer* remoto richiede l'identificativo dell'utente e la *password*.

Una volta attuata la fase di *login* si può finalmente disporre delle risorse presenti sul sistema remoto.

L'unica funzionalità importante di **Internet** che i principali *browser* non integrano nella loro interfaccia è proprio il telnet. Ciò significa che non è possibile visualizzare direttamente dall'interno di questi *browser* schermate in emulazione terminale.

## Gopher

Il *gopher* (ormai quasi totalmente inutilizzato) è essenzialmente uno strumento di organizzazione testuale dell'informazione presente su **Internet**. Il principio su cui si fonda questa organizzazione è quello della struttura gerarchica.

La struttura gerarchica è un sistema di organizzazione delle risorse molto efficiente. È molto diffuso in ambito informatico. Esempi tipici di strutture gerarchiche sono l'organigramma di un'azienda, l'albero genealogico, i sistemi di classificazione bibliotecaria.

Per consultare le informazioni organizzate nel *gopherspace* (o almeno ciò che ne resta) bisogna naturalmente avere accesso ad un *client gopher*. Ne sono stati sviluppati per tutti i sistemi operativi, sia in ambienti a caratteri sia in ambienti dotati di interfaccia grafica, ma poiché i *browser Web* sin dall'inizio hanno integrato il supporto al protocollo *gopher*, la maggior parte è ormai stata abbandonata.

Un modo alternativo (ed assai più comodo) per accedere ai residui *gopher server* presenti su **Internet**, come detto, è costituito dall'uso di un comune *browser Web*, come *Netscape* o *Internet Explorer*.

L'aspetto di un *gopher server* nella finestra di un *browser* è sostanzialmente uguale a quello di un *ftp server*: un elenco di voci affiancate da piccole icone che ne denotano la tipologia.

## Servizi Multimediali

I documenti multimediali interattivi integrano diverse modalità comunicative e

consentono l'interazione con l'utente.

I documenti multimediali interattivi più completi contengono audio, video e ambienti interattivi in grafica 3D (tridimensionale).

Il formato principale dei documenti del servizio *Web* è html, che però ha limitatissime capacità grafiche, video, audio ed interattive.

Per avere documenti interattivi un'alternativa importante che non approfondiamo è l'uso di un linguaggio di programmazione: *Java* e *JavaScript* sono due esempi di linguaggi capaci di animare i documenti.

Diversamente da html, i linguaggi di programmazione non sono alla portata di tutti gli autori di documenti **Internet**.

Al momento esistono pochi *standard* de-jure per servizi multimediali; quindi, le applicazioni che per prime hanno portato sulla rete un certo tipo di formato di documento multimediale sono diventate *standard* di fatto.

Ecco alcuni esempi di *software* per documenti multimediali:

- *Quicktime*: [www.apple.com/quicktime](http://www.apple.com/quicktime)
- *Flash* e *Shockwave*: [www.macromedia.com](http://www.macromedia.com)
- *RealPlayer*: [www.realplayer.com](http://www.realplayer.com)

## Formati audio

L'informazione audio viene codificata in digitale in moltissimi modi: i formati più comuni sono .wav, .mid, .mp3, .ra

- .wav, creato da *IBM* e *Microsoft*, rappresenta onde sonore in formato digitale di qualità variabile; infatti, la dimensione di un .wav dipende dalla qualità di digitalizzazione (1 minuto in qualità CD = 10 MB circa);
- .mid denota il formato MIDI, che contiene istruzioni eseguibili da strumenti musicali elettronici; quindi i *file* .mid sono molto compatti (un minuto = circa 10 KB);
- .mp3 è un formato audio digitale compresso di qualità quasi CD (un minuto in qualità quasi CD = circa 1 MB);
- .ra è un formato audio ([www.realaudio.com](http://www.realaudio.com)) per trasmissione **Internet** a banda stretta, fornito in varie versioni in funzione della connessione modem disponibile.
- .aif, creato da *Apple*, sta per *Audio Interchange File format*.

Il programma riproduttore audio più comune su *Windows* si chiama *Winamp* ([www.winamp.com](http://www.winamp.com)).

Un programma su *Windows* molto versatile per effettuare conversioni audio è *AudioCatalyst* ([www.xingtech.com](http://www.xingtech.com)).

È ad esempio molto facile convertire un brano in versione CD musicale nel formato mp3.

## Servizi audio di rete

Distinguiamo due principali classi di servizi: *broadcast* (radio digitale) e punto a punto (*unicast*: telefono su IP)

- **Broadcast: Digital radio.** È un servizio che permette di ascoltare la radio usando un *computer* collegato a **Internet**. Ad esempio si veda il sito [www.harmony-central.com](http://www.harmony-central.com) per alcuni programmi audio.
- **Unicast: telephone VoIP (Voice over Internet telephone).** Con *VoIP* la voce di chi utilizza il servizio viene digitalizzata e spedita via *TCP/IP* (attraverso **Internet**). Al momento esistono tre tipi di *VoIP*: *Pc to Pc*, *Pc to Phone*, *Phone to Phone*, che assicurano rispettivamente la comunicazione tra due *computer*, tra un *computer* e un telefono di rete fissa o mobile e tra due telefoni.

Esempi:

- [www.net2phone.com](http://www.net2phone.com)
- [www.voispring.it](http://www.voispring.it)
- [www.picus.it](http://www.picus.it)
- [www.chocophone.com](http://www.chocophone.com)

## Formati grafici

Anche i formati grafici sono tantissimi e permettono di rappresentare immagini in formato digitale. I principali in uso su **Internet** sono .gif e .jpg, entrambi basati su grafica bitmap (l'immagine è rappresentata come una matrice di *pixel* = *picture element*).

- **GIF:** formato grafico su 256 colori, compresso senza perdita di informazione, eccellente per icone o logo ovvero immagini che non hanno bisogno di alta definizione.
- **JPEG:** formato grafico su milioni di colori, compresso con perdita di informazione, eccellente per foto (o immagini che necessitano di un certo grado di definizione).
- **Animazione GIF:** visualizzazione di immagini GIF in rapida successione che serve per rappresentare semplici animazioni con bassa definizione.

Le immagini GIF usano la compressione *lossless* (senza perdita di informazione) LZW, che cerca le sequenze di *pixel* ripetute: dunque questo formato è adatto ad immagini in bianco e nero, specie se contrastate, ma non è adatto a comprimere fotografie.

**Nota:** il formato JPEG per molte immagini è più compatto di GIF. Esempio: foto di persona, .jpg 156KB, .gif 416KB.

Più importante è il fatto che siccome la compressione JPEG è *lossy* (con perdita di informazione/definizione) si può decidere quanta informazione verrà eliminata (a seconda della definizione desiderata).

## Comprimere documenti video

Problema: un'immagine video digitale in risoluzione 640x480 con un milione di colori occupa circa 1MB. La qualità video cinematografica vuole 24 fotogrammi al secondo, quindi un film di due ore richiederebbe 172.800 MB, pari a circa 260 CD audio.

La grande diffusione di connessioni *Internet* a banda stretta ha fatto nascere formati video speciali: il formato *RealVideo* ad esempio garantisce da 1 a 6 fotogrammi al secondo, risoluzione 80x120, con audio buono (stereo, 16 *bit*, 44KHz).

Nel 1988 nasce il gruppo di lavoro MPEG (*Moving Picture Expert Group*) per mettere a punto un sistema di codifica (CODEC) di video e audio compressi *lossy*.

- Mpeg-1 (ISO 11172, 1991) riguarda sia audio stereo che video (qualità simile a VHS); è uno *standard* che prevede tre livelli di compressione, per 30 fotogrammi/sec, risoluzione tipica 352x240 (max 512x324), audio CD.
- Mpeg-2 (estensione .mpg, ISO 13818, 1997) riguarda audio multicanale per *Home Cinema* (quindi è usato per DVD) campionato a 24KHz. Ha 60 fotogrammi/sec, risoluzione 1280x720, audio CD. **Nota:** Mpeg-2 layer 3 è oggi noto come MP3 (estensione .mp3)
- Mpeg-4 (.avi) è la specifica di *standard* più moderna, ed al momento si incarna in una tecnologia chiamata DivX: un intero film (110 minuti, in media) sta su un CDROM

## Formati di streaming

La trasmissione audio-video su *Internet* a banda stretta si basa su *streaming*, ovvero la riproduzione di audio e/o video contemporaneamente al loro *download*; le principali piattaforme *streaming* sono al momento *RealNetworks*, *Microsoft* e *Apple*.

*RealNetworks* (www.real.com) pubblicò il *RealPlayer* (estensione .rm) nel 1995, nel 2000 è alla versione 8.

*Microsoft* include il *Media Player* (versione 7 nel 2000) con tutte le versioni di *Windows* (WindowsMedia.com) ed ha circa 2/3 degli utenti di *RealMedia*.

*Apple* (www.apple.com/quicktime) ha sviluppato il formato *QuickTime* (estensione .mov) che nel 1999 è alla versione 4. La sua forza sta nel fatto di essere davvero multipiattaforma; inoltre supporta parecchi tipi di contenuto: testo, immagini ferme, 3-D, e *Virtual Reality*.

## Servizi basati su streaming (WebCasting)

Il *WebCasting* è la diffusione sistematica di informazioni video in formati *streaming*.

Un sito indice dei canali *WebCasting* è www.mediachannel.com.

Le principali TV del mondo hanno siti *Web* in cui propongono filmati da vedere via *streaming*.





sito Internet della CNN

## Formati per documenti interattivi

I documenti interattivi sono documenti che includono animazioni e suoni e sono capaci di interagire con l'utente (come i videogiochi). Si usano tipicamente per presentazioni graficamente complesse, specie nel campo dell'*e-learning*.

*Macromedia Director* è il più famoso programma di *authoring* di documenti di questo tipo; *Shockwave* è un *plug-in* per *browser* che consente di riprodurre documenti realizzati con *Macromedia Director*. A partire dalla versione 6.0 di *Director*, *Shockwave* permette lo *streaming*.

Nel 1997 è apparso *Flash 2.0*, allo scopo di portare su *Web* alcune caratteristiche importanti di *Director*.

La forza di *Flash* è nel formato vettoriale con cui opera. Consente di ottenere filmati (*movie*) compatti e di ottima qualità grafica, che possono zoomare a tutto schermo senza perdita di qualità. Un altro vantaggio delle immagini vettoriali è che il filmato può racchiudere informazioni aggiuntive. Ad esempio, agli oggetti possono essere associate delle proprietà particolari, come l'attivarsi ad un click del *mouse*.

Perché dunque i *browser* non visualizzano immagini vettoriali in luogo delle GIF e delle JPG (che sono *bitmap*)? La spiegazione è che non esiste uno *standard* industriale per la trasmissione di immagini vettoriali.

Al momento (2002) la partita degli *standard* si gioca tra *Adobe* e *Macromedia*: la prima ha sottoposto al W3C un formato vettoriale basato sul PDF chiamato PGML, mentre la seconda, con *Flash*, ha imposto de facto il proprio *standard*.

## World Wide Web (WWW)

Il *World Wide Web* (WWW o più semplicemente *Web*), che letteralmente significa ragnatela intorno (grande quanto il) al mondo, è uno dei tanti servizi disponibili su *Internet*. Anche se il *Web* senza dubbio è uno dei servizi più usati dagli utenti (forse il più usato in assoluto) desideriamo puntualizzare che il *Web* non coincide con *Internet*, ma è uno dei tanti servizi a cui è possibile accedere collegandosi a *Internet*.

In estrema sintesi il *Web* è una collezione di documenti ipertestuali e multimediali contenuti su alcune macchine (*Web server*) collegate a *Internet*. Tramite applicazioni specifiche chiamate *browser* (clienti *Web*) è possibile navigare in questo mare di documenti, visualizzarli e memorizzarli sul proprio calcolatore.

Il *Web* è un vero e proprio sistema di editoria elettronica *on-line*.

Uno degli aspetti del *Web* che ha maggiormente contribuito alla sua diffusione è senza dubbio la capacità di gestione e controllo dei documenti multimediali, e dunque dei linguaggi utilizzati per la loro creazione. Un ruolo importante nel processo di espansione e diffusione del *Web* è rivestito dalle grandi aziende produttrici di *browser*. Nel corso degli anni tanto *Microsoft* quanto *Netscape*, man mano che nuove versioni dei loro *browser* venivano sviluppate, hanno introdotto innovazioni ed estensioni. Questa corsa all'ultima innovazione, se molto ha migliorato l'aspetto e la fruibilità delle pagine pubblicate su *Web*, ha avuto degli effetti deleteri sul piano della portabilità dei documenti. Per ovviare al rischio di una babele telematica, ed evitare che le tensioni indotte dal mercato limitassero l'universalità di accesso all'informazione *on-line*, è stato costituito il *World Wide Web Consortium* (W3C). Si tratta di una organizzazione non profit ufficialmente deputata allo sviluppo degli *standard* tecnologici per il *Web* che raccoglie centinaia di aziende, enti e centri di ricerca coinvolti più o meno direttamente nel settore. In questi ultimi anni il W3C ha prodotto una serie di specifiche divenute, o in procinto di divenire, *standard* ufficiali su *Internet*. Tutti i materiali prodotti dal W3C sono di pubblico dominio, e vengono pubblicati sul sito *Web* del consorzio.

## Brevi cenni storici sul Web

- Nel 1945, *Vannevar Bush* (tecnologo del MIT e consulente del governo USA) scrive un articolo in cui descrive la nozione di ipertesto. ([www.theatlantic.com/unbound/flashbks/computer/bushf.htm](http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm))
- Nel 1965 *Ted Nelson* inventa la parola **ipertesto** nell'articolo: *A File Structure for the Complex, the Changing, and the Indeterminate 20th National ACM Conference, New York*. Vedi [www.sfc.keio.ac.jp/~ted/TN/PUBS/LM/LMpage.html](http://www.sfc.keio.ac.jp/~ted/TN/PUBS/LM/LMpage.html)
- Nel 1968 *Doug Engelbart* costruisce il prototipo di un *onLine System* (NLS) con funzioni di *editing* ipertestuale, e inventa il *mouse* per interagire con tale sistema. Vedi [stanford.edu/MouseSite/](http://stanford.edu/MouseSite/)
- Nel 1989, *Berners-Lee* del CERN specifica un sistema di gestione di documenti ipertestuali distribuiti.

- Alla fine del 1990 viene realizzato un prototipo.
- Nel febbraio 1993 venne rilasciato *Mosaic*, un *browser* scritto dallo studente *M. Andreessen*.
- Alla fine del 1994 *Andreessen* fonda *Netscape Corporation* (nel 1999 verrà acquisita da AOL, *America On Line*).
- Nel dicembre 1994 nasce il W3C.

## Dimensioni del www

Riportiamo alcune cifre che danno l'idea delle dimensioni del *Web* e della sua continua vertiginosa espansione.

All'inizio dell'anno 2000 c'erano *on-line* un miliardo di pagine circa. Alla fine del 2000 c'erano *on-line* 2.5 miliardi di pagine circa. Alla fine del 2001 4 miliardi di pagine circa. Con una crescita *media* di più di 7 milioni di pagine al giorno.

Alla fine del 2001, il motore di ricerca *Google* classifica 3 miliardi di pagine, inclusi *newsgroup* e immagini.

Il mondo produce annualmente da uno a due *exabyte* (10<sup>18</sup> *byte*) di dati, ovvero all'incirca 250 *megabyte* per ogni abitante della terra.

I documenti stampati, di qualsiasi genere, sono solo lo 0.003% del totale. Le informazioni memorizzate su supporto magnetico sono di gran lunga le più numerose, e crescono rapidamente: ogni anno raddoppiano.

## Concetti Logici di base: ipertesti e multimedialità

Il *Web* in estrema sintesi è un gigantesco ipertesto multimediale distribuito.

Il concetto di multimedialità si riferisce alla tipologia dell'informazione rappresentata, mentre il concetto di ipertesto si riferisce alla struttura ed alla organizzazione dell'informazione stessa.

### Ipertesto

Dal punto di vista logico un ipertesto è un sistema di organizzazione delle informazioni (testuali, ma non solo) in una struttura non sequenziale, bensì reticolare.

Ciò significa che un ipertesto permette di saltare da un punto ad un altro del documento o addirittura saltare da un documento ad un altro seguendo *link* specifici messi a disposizione dall'autore del documento.

Dal punto di vista della implementazione concreta, un ipertesto digitale si presenta come un documento elettronico in cui alcune porzioni di testo o immagini presenti sullo schermo, evidenziate attraverso artifici grafici (icone, colore, tipo e stile del carattere), rappresentano i diversi collegamenti disponibili nella pagina. Questi funzionano come dei pulsanti che attivano il collegamento e consentono di passare, sullo schermo, al documento di destinazione. Il pulsante viene premuto attraverso un dispositivo di *input*, generalmente il *mouse*, una combinazione di tasti, o un tocco su uno schermo *touch-screen*. L'altro aspetto che fa dell'ipertesto elettronico uno strumento



comunicativo dalle enormi potenzialità è la interattività che esso consente al fruitore, non più relegato nella posizione di destinatario più o meno passivo del messaggio, ma capace di guidare e indirizzare consapevolmente il suo atto di lettura.

## Multimedia

Con multimedialità ci si riferisce alla possibilità di inserire contemporaneamente in uno stesso documento informazione di tipo diverso. Ad esempio, immagini, suoni e filmati. Questa informazione ovviamente viene rappresentata in forma digitale appropriata usando *standard*, linguaggi e strumenti tra loro molto diversi.

## Distribuito

Un ipertesto multimediale può essere spezzato in più parti (memorizzate in *file*). Ognuna di queste parti è tipicamente memorizzata su un *host* di rete. Se un documento è memorizzato su più *host* viene detto distribuito.

L'incontro tra ipertesto, multimedialità e interattività rappresenta dunque la nuova frontiera delle tecnologie comunicative. Il problema della comprensione teorica e del pieno sfruttamento delle enormi potenzialità di tali strumenti, specialmente in campo didattico, pedagogico e divulgativo è naturalmente ancora in gran parte aperto: si tratta di un settore nel quale vi sono state negli ultimi anni, ed è legittimo aspettarsi negli anni a venire, innovazioni di notevole portata.

## Come funziona il Web - Architettura client-server.

La navigazione in **Internet** si basa su interazione di tipo *client-server*. Le funzioni elaborative necessarie per rendere possibile la navigazione in **Internet** sono distribuite tra *client* e *server* in modo da distribuire il carico di lavoro e ottimizzare l'efficienza complessiva del sistema.

## Client

Per navigare in **Internet** è necessario possedere un *software* (applicativo) chiamato *browser* (esempi: *Netscape Communicator*, *MS Internet Explorer*). Il *browser* costituisce lo strumento di interfaccia tra l'utente e il sistema *Web* e si occupa sostanzialmente di:

- ricevere ed eseguire i comandi dell'utente;
- richiedere ad un *server* i documenti;
- interpretare i formati di codifica degli oggetti che costituiscono ogni singolo documento e presentarli all'utente sul suo monitor (fase di *rendering*);
- invocare applicazioni esterne (*helper*);
- riprogrammarsi mediante *plug-in*;
- eseguire codice *JavaScript* o *Java*.

## Server

Un *server Web*, o più precisamente *server http* (vedremo poi il perché di questo nome)

si occupa di memorizzare, reperire e inviare i documenti (e tutti gli oggetti digitali che li costituiscono) richiesti dai *client*. Nel momento in cui il *client* (utente) invia una richiesta http al *server* opportuno, con l'indicazione del documento che vuole ricevere, il *server* interpreta la richiesta e invia gli oggetti che compongono il documento richiesto al *client*.

## Protocolli del Web

Il *client* e il *server* comunicano e rendono possibile la navigazione. Per comunicare utilizzano protocolli specificatamente progettati a tal fine.

Nel seguito daremo una descrizione sintetica dei meccanismi e dei protocolli principali che permettono di navigare in *Internet*.

Il *Web* è un servizio *Internet* che si basa su tre protocolli fondamentali:

### Uniform Resource Locators (url)

Schema di denominazione di risorse in *Internet* (tipicamente *file* in formato html). Quando parliamo di un url ci riferiamo al nome completo ed univoco che una certa risorsa ha in *Internet*. Ad esempio diciamo che [www.cs.unibo.it](http://www.cs.unibo.it) è l'url del sito *Internet* del dipartimento di Scienze dell'Informazione dell'università di Bologna.

### HyperText Transfer Protocol (http)

Http è il protocollo di livello applicazione che regola il trasferimento dei *file* html in *Internet*.

- I *server* http distribuiscono documenti ipertestuali.
  - I *clienti* http (*browser*) li visualizzano.
- Server* e *client* per comunicare fanno riferimento al protocollo http.

### HyperText Markup Language (html) per ipertesti

HTML è un linguaggio di *markup*, ossia un linguaggio che permette di specificare come un certo testo deve essere formattato (colori, collegamenti ipertestuali, sfondi e molte altre caratteristiche grafiche) al momento della sua visualizzazione sul nostro *display*.

## Siti Web

Un sito *Web* è definito da:

- un indirizzo, o url (esempio <http://www.cs.unibo.it>) al quale corrisponde un *server* http;
- un insieme di documenti (scritti di solito in html) sotto il controllo del *server* http;
- un insieme di servizi, che è variabile; nel caso la parte di servizi sia preponderante su quella dei documenti il sito si definisce portale.

## Come funziona http

Ogni sito *Web* (*server* http corrispondente) ha un processo servente in ascolto sulla

porta 80 in attesa di connessioni http; dopo che è stata stabilita una connessione, il cliente invia una richiesta e il server invia una risposta; quindi la connessione viene abbandonata (http è un servizio non orientato alla sessione).

Esempio: supponiamo di voler seguire il link:

[www.w3.org/hypertext/WWW/TheProject.html](http://www.w3.org/hypertext/WWW/TheProject.html).

Le fasi che si succedono tra il click dell'utente e la visualizzazione della pagina sono le seguenti:

- Il *browser* determina l'url (controlla ciò che è stato selezionato).
- Il *browser* chiede al dns l'indirizzo di [www.w3.org](http://www.w3.org).
- Il dns risponde con 18.23.0.23 (indirizzo fisico della macchina [www.w3.org](http://www.w3.org)).
- Il *browser* apre una connessione *tcp* sulla porta 80 della macchina con indirizzo 18.23.0.23.
- Quindi invia un comando *GET* /hypertext/WWW/TheProject.html.
- Il *server* [www.w3.org](http://www.w3.org) invia il *file* TheProject.html.
- Viene rilasciata la connessione *TCP*.
- Il *browser* ripete il ciclo per ciascuna immagine (o componente di altro tipo) in TheProject.html.

Ogni interazione http è una richiesta ASCII, seguita da una risposta *mime*.

Http è un protocollo *client-server*, progettato per la distribuzione rapida di documenti ipertestuali su reti *TCP/IP*.

È un protocollo *connection-less*, che significa che ogni richiesta di un cliente viene elaborata dal *server* indipendentemente: il *server* non ha memoria della sessione.

Esempio: se un *server* http deve trasmettere un documento html contenente 10 immagini, occorrono 11 connessioni.

Ogni risorsa è specificata da una url.

**Nota.** Sebbene http sia stato inventato per il *Web*, è più generale del necessario, con funzioni pensate per future applicazioni.

## Publicare pagine sul Web

In estrema sintesi, per pubblicare un documento su *Web* occorre:

- poter usare una url pubblica di un *server Web*;
- poter scrivere in un disco sotto il controllo del *server Web* (ovvero fare *upload* via ftp);
- scrivere il documento direttamente in html oppure usare un *editor* html (esempio *DreamWeaver*);
- preparare, linkare e fare l'*upload* sul *server* delle figure (se presenti);
- fare l'*upload* del documento sul disco del *server*.

## Helper e plug-in

Il *browser* può invocare l'aiuto di altre applicazioni per gestire documenti in formati particolari. Quando un *browser* viene progettato non è possibile prevedere quali nuovi formati di *file* e quali tecnologie per la loro gestione saranno introdotti in futuro. Inoltre i *browser* sono applicazioni molto complesse e di grandi dimensioni. Come conseguenza di queste considerazioni i *browser* si appoggiano e usano altre applicazioni per gestire formati di *file* specifici. Queste applicazioni sono sostanzialmente di due tipi: *helper* e *plug-in*.

### Helper

Un *helper* è un programma indipendente dal *browser*, ma da esso invocabile direttamente per gestire la visualizzazione di particolari tipi di documenti. Esempi di *helper*:

- *GhostView* per documenti in formato *PostScript*.
- *RealPlayer* per documenti *streaming* in formato *Realaudio*.

### Plug-in

Un *plug-in* è un componente *software* del *browser* in grado di arricchirne la capacità di gestire documenti in formati speciali. Esempi di *plug-in*:

- *Acrobat plug-in* per documenti in formato pdf.
- *Quicktime plug-in* per filmati in vari formati.

I *plug-in* di solito sono programmi disponibili in rete, caricati automaticamente dal *browser* quando si trova di fronte ad un documento in formato speciale sconosciuto.

La sostanziale differenza tra *helper* e *plug-in* è che l'*helper* è un programma esterno al *browser* che viene attivato al momento di gestire *file* di un certo formato, mentre il *plug-in* è un modulo *software* che si può aggiungere al *browser* e che viene usato al suo interno senza dover attivare nessun altro programma oltre al *browser* stesso.

*Netscape 2.0* è stato il primo *browser* che ha supportato l'idea tecnologica secondo la quale case produttrici diverse da quella del *browser* offrono componenti *software* che lo integrano.

Un buon sito in cui si trovano molti *plug-in* è:

[browserwatch.internet.com/plug-in/plug-in-win.html](http://browserwatch.internet.com/plug-in/plug-in-win.html)

## Html

Un documento html è un documento scritto nel linguaggio di *markup* html, che contiene i *tag* (delimitatori sintattici) necessari ad informare il *browser* sulle modalità di formattazione.

**Nota:** Un documento contenente testo senza alcun *tag* html viene comunque visualizzato dal *browser*, ma senza alcuna formattazione.

# Usare Internet

Luciano Margara

## Url

Il sistema delle url (*Uniform Resource Locator*) è uno schema di denominazione per specificare come e dove sia possibile reperire una risorsa **Internet** (di solito un *file*). Esempio di url:

`http://www.cs.unibo.it/~margara/index.html`

Le risorse sono di solito *file* in formato html (ma non solo) contenute in un *server* http.

## Esempi di URL valide

`http://www.cs.unibo.it/~margara/lucidi.html`

`https://pyconv.cs.unibo.it`

`ftp://cs.unibo.it/pub/margara/`

`gopher://gopher.di.unipi.it/`

`telnet://129.135.2.6:4801`

`file:///C:/didattica/lucidi.htm`

Un url è composto da due parti. La prima parte specifica come accedere alla risorsa. La seconda parte specifica dove si trova la risorsa.

## Esempio 1

`http://www.cs.unibo.it/~margara/lucidi.html`

La parte `http:` indica che per accedere alla risorsa (specificata nella seconda parte dell'url) useremo il protocollo `http` e che la risorsa è contenuta in un *server Web* (o equivalentemente *server http*).

La parte `www.cs.unibo.it` dell'url denota il nome logico del *Web server* che contiene la risorsa.

La parte `/~margara/lucidi.html` specifica il nome del *file* html al quale vogliamo accedere (che vogliamo visualizzare). Per poter accedere a un *file* html (o di altro tipo) memorizzato su una macchina remota (*server http*) dobbiamo conoscere come il *file system* di quella macchina assegna i nomi ai *file*. Secondo le regole del *file system* della macchina `www.cs.unibo.it` il nome `/~margara/lucidi.html` individua un *file* html contenuto nella *directory* `margara` che è una *directory* contenuta in una zona specifica (la zona dedicata alle *home directory* dei docenti del dipartimento di Scienze dell'Informazione) del *file system* specificata dal carattere `~`.

## Esempio 2

<https://pyconv.cs.unibo.it>

La parte `https:` indica che per accedere alla risorsa useremo il protocollo `https` che è la versione sicura del protocollo `http`. Sicura significa che durante la connessione al *Web server* specificato nella seconda parte dell'url verranno usati accorgimenti specifici per rendere sicura la connessione (tipicamente meccanismi basati su algoritmi di crittografia).

La parte `pyconv.cs.unibo.it` dell'url denota il nome logico del *web server* che contiene la risorsa.

Quando l'url termina senza specificare il nome della risorsa (del *file*) automaticamente si accede ad una risorsa chiamata risorsa di *default*. Se non diversamente specificato tale risorsa è il *file* di nome `index.html`.

## Esempio 3

`telnet://129.135.2.6:4801`

In questo esempio la prima parte dell'url ci dice che stiamo usando il protocollo `telnet` per accedere alla macchina con indirizzo *IP* `129.135.2.6` attivando un collegamento sulla porta fisica numero `4801`.

## Posta Elettronica

Il servizio di posta elettronica (*e-mail*) consente di scambiarsi sia messaggi di testo sia, utilizzando gli opportuni strumenti, ogni altro tipo di *file*.

La comunicazione tramite posta elettronica è asincrona ovvero non è necessario che il destinatario del messaggio di posta elettronica sia collegato a *Internet* nel momento in cui il mittente gli invia il messaggio.

Ogni messaggio ha un mittente e uno o più destinatari.

Per poter spedire o ricevere un messaggio di posta elettronica è necessario possedere un indirizzo di posta elettronica (*e-mail address*).

L'indirizzo solitamente ci è assegnato dal nostro fornitore di connettività (*Internet Service Provider*, *ISP*) oppure ci viene offerto da un altro fornitore di servizi di rete come ad esempio `hotmail.com`. I dipendenti di università o di enti pubblici solitamente ricevono un indirizzo di posta elettronica dagli amministratori di rete della propria organizzazione.

Ad ogni indirizzo di posta elettronica corrisponde una casella postale (*mailbox*) che è memorizzata sull'*hard disk* della macchina (*mail server*) di proprietà del fornitore di servizi che ci ha assegnato l'indirizzo di posta. Nella casella postale vengono depositati automaticamente i messaggi di posta indirizzati al proprietario della casella.

Questo significa che non c'è bisogno che il nostro *computer* sia sempre collegato ad *Internet*, in attesa dei messaggi che ci potrebbero arrivare: è il *computer* (*mail server*) di chi ci ha fornito l'indirizzo di posta elettronica che si assume questo incarico per noi. Quando decideremo di collegarci a *Internet* controlleremo nella nostra casella postale

se ci sono messaggi in attesa.

Il *mail server* invece è un calcolatore collegato a **Internet** 24 ore su 24. Sul *mail server* gira un programma chiamato demone di posta che ha un solo scopo: quello di ricevere messaggi di posta indirizzati a una delle caselle postali ospitate dal *server* in questione.

## Indirizzi di posta elettronica

La forma generale di un indirizzo è la seguente:

nomeutente@nomecomputer

La parte di indirizzo alla sinistra del simbolo @ (detto chiocciola o in inglese *at*) identifica l'utente in maniera univoca all'interno del *mail server* che ospita la sua *mailbox*. Spesso si tratterà del nostro cognome, o di un codice, o di un nomignolo che ci siamo scelti. L'importante è che non ci siano due utilizzatori di quel sistema con lo stesso identificativo.

La parte di indirizzo a destra del simbolo @ identifica invece in maniera univoca, all'interno dell'intera rete **Internet**, il particolare sistema informatico presso il quale l'utente è ospitato, e corrisponde all'indirizzo simbolico dell'*host*. L'indirizzo di un utente denominato margara collegato all'*host* denominato cs.unibo.it sarà dunque:

margara@cs.unibo.it

## Come viene trasmessa la posta

La posta elettronica viaggia su *Internet* dal *mail server* che ospita la casella postale del mittente al *mail server* che ospita la casella postale del destinatario. Questo tragitto viene compiuto seguendo le regole descritte nel protocollo SMTP (*Simple Mail Transfer Protocol*) che è per l'appunto il protocollo di livello applicazione che gestisce il trasferimento della posta elettronica.

Per editare un messaggio di posta è necessario utilizzare un programma chiamato cliente di posta (esempio: *Outlook*, *Eudora*). Una volta che il messaggio è stato scritto è necessario trasferirlo dal nostro *computer* al *mail server* (che poi si occuperà di inviarlo fisicamente tramite **Internet** verso il destinatario). Anche questo tragitto viene compiuto seguendo le regole descritte nel protocollo SMTP.

Il cliente di posta permette anche di leggere la posta elettronica che giace nella nostra casella sul *mail server*.

Una volta collegati a **Internet** è infatti possibile chiedere al cliente di posta (che gira sulla nostra macchina) di trasferire la posta a noi indirizzata dal *mail server* al nostro calcolatore. Questo tragitto viene compiuto seguendo le regole descritte nel protocollo POP (*Post Office Protocol*) o nel protocollo IMAP (*Interactive Mail Access Protocol*).

La fase di lettura e scrittura dei messaggi può quindi essere eseguita in modalità *off-line*, ovvero senza essere connessi a **Internet**

## Formato di un messaggio



Ecco un esempio di messaggio di posta elettronica:

From: Fabio Vitali <fabio@CS.UniBO.IT> Date: Mer gen 15, 2003 17:25:47 Europe/Rome To: Luciano Margara <margara@CS.UniBO.IT> Subject: Prova Received: by le (mbox margara) (with Cubic Circle's cucipop (v1.31 1998/05/13) Wed Jan 15 17:25:55 2003) Received: from [130.136.2.220] (genesis.cs.unibo.it [130.136.2.220]) by CS.UniBO.IT (8.9.3/8.9.3/Debian 8.9.3-6) with ESMTP id RAA29182 for <margara@cs.unibo.it>; Wed, 15 Jan 2003 17:25:45 +0100 X-From\_: fabio@CS.UniBO.IT Wed Jan 15 17:25:46 2003 User-Agent: Microsoft-Entourage/10.0.0.1309 Message-Id: <BA4B4A1B.D4BE%fabio@cs.unibo.it> Mime-Version: 1.0 Content-Type: text/plain; charset=US-ASCII Content-Transfer-Encoding: 7bit Questa e' una prova

Un messaggio (secondo lo *standard* RFC 822) include alcuni campi di intestazione (qui sotto elencati con il relativo significato), una linea bianca e quindi il corpo del messaggio.

Campo	Significato
<i>To:</i>	Indirizzo per destinatario primario
<i>Cc:</i>	Indirizzo per destinatario secondario
<i>Bcc:</i>	Indirizzo invisibile agli altri ( <i>blind cc</i> )
<i>From:</i>	Persona che ha creato il messaggio
<i>Sender:</i>	Indirizzo elettronico del vero mittente
<i>Received:</i>	Ogni <i>mail transfer agent</i> sul percorso mittente-destinatario si identifica
<i>Return-Path:</i>	Individua un percorso verso il mittente
<i>Date:</i>	Data e ora in cui il messaggio è stato inviato
<i>Reply-To:</i>	Indirizzo al quale inviare le risposte
<i>Message-Id:</i>	Identificativo del messaggio
<i>In-Reply-To:</i>	<i>Message-Id</i> del messaggio cui si sta rispondendo
<i>References:</i>	Altri <i>Message-Id</i> importanti
<i>Keywords:</i>	Parole chiave scelte dall'utente
<i>Subject:</i>	Sommario del messaggio (su una sola linea)

## Allegati (attachments)

I clienti di posta elettronica dispongono di un'opzione che serve ad allegare al messaggio uno o più documenti multimediali in forma di allegati (*file*).

Un allegato (*attachment*) è un *file* incluso in forma compressa in un messaggio di *email*.

L'*attachment* di solito viene compresso automaticamente dal cliente di posta elettronica prima di spedirlo assieme al messaggio.

Al momento di ricevere un messaggio con allegato, il destinatario potrà vederlo o editarlo, ma solo se possiede una applicazione in grado di gestire il formato



dell'allegato stesso.

## Gestione automatica della posta

Uno strumento prezioso per molti utenti è la capacità di impostare dei filtri. Un filtro è un insieme di regole che vengono verificate dal cliente di posta quando arriva un messaggio.

Una regola potrebbe stabilire per esempio che qualsiasi messaggio proveniente da `margara@cs.unibo.it` va salvato in un *file* particolare, oppure visualizzato con un *font* a 24 punti lampeggiante in grassetto di colore rosso (oppure va eliminato automaticamente!)

## Mailing list

Una *mailing list* è un servizio di comunicazione uno-a-molti che si basa sulla posta elettronica.

Esempio: *A.Word.A.Day* ([www.wordsmith.org/awad](http://www.wordsmith.org/awad)) è una *mailing list* (circa 100.000 persone in più di 120 nazioni) che spedisce un messaggio al giorno a tutti coloro che si sono iscritti alla *mailing list*. Il messaggio spiega il significato e l'uso di una parola inglese.

Le *mailing list* possono essere gestite:

- manualmente, da un moderatore che mantiene una lista di indirizzi di posta elettronica (di solito lato *client*);
- automaticamente, da un servizio che accetta iscrizioni via posta elettronica.

**Esempio:**

per iscriversi alla *mailing list* **Internet** *Tourbus*, occorre mandare un messaggio *email* a `listserv@listserv.aol.com`, includendo nel corpo del messaggio il comando:

```
subscribe tourbus Paolo Rossi
```

## Netiquette (regole di galateo) per e-mail

Alcune regole che è bene rispettare usando la posta elettronica:

- Quando si accede **Internet** da un ufficio, verificare con il datore di lavoro la questione della proprietà dei messaggi, perché le leggi al riguardo variano da paese a paese.
- A meno che non si usi un filtro di criptazione, conviene assumere che la posta su **Internet** non sia mai sicura.
- Rispettare il *copyright* sui materiali riprodotti (anche se ogni paese ha una propria legislazione sul *copyright*).
- In caso di inoltro di un messaggio ricevuto, non modificarne il testo. Se si tratta di un messaggio personale e lo si vuole re-inviare ad un gruppo, occorre chiedere preventiva autorizzazione all'autore originale del messaggio.

- Non inviare mai per posta elettronica lettere a catena di Sant'Antonio. Se ne ricevete una, avvertire il sistemista.
- Si sia rigorosi in quel che si spedisce e tolleranti in quel che si riceve. Non inviare messaggi aggressivi (*flames*) anche se provocati. Può capitare di diventare vittime di qualche *flame*, e in tal caso è prudente non rispondere.
- È bene rendere le cose facili per chi riceve. Siccome molti *mailer* eliminano lo *header*, compreso l'indirizzo del mittente, assicurarsi di includere nella firma a fine messaggio i dati rilevanti: Non mandare messaggi anonimi.
- Usare maiuscole e minuscole. SE SI USANO SOLO LE MAIUSCOLE È COME SE SI STESSE URLANDO.

## Comunicare in Rete: talk, IRC

Un sistema di *chat* è un servizio di comunicazione sincrona (i partner della comunicazione devono essere collegati a *Internet* e pronti a comunicare simultaneamente).

### **Talk**

Un programma di *talk* (*Unix: talk; Windows: Wintalk*) permette di aprire una sessione con un utente remoto per effettuare una conversazione in forma testuale; entrambi i corrispondenti devono possedere un cliente specifico ossia un programma che gira localmente sulla macchina di chi desidera attivare la sessione di *talk* e che permette di inviare e ricevere testo.

Nota: alcune società offrono servizi di *talk* capaci di usare come cliente una normale cornetta telefonica; in pratica ciò vuol dire che si può usare *Internet* per telefonare.

### **IRC (*Internet Relay Chat*)**

IRC è un protocollo (ma anche un'applicazione) che permette il dialogo di più persone mediante messaggi testuali.

Per usare il servizio occorrono clienti IRC che si connettono a *server* specifici. Quando ci si connette ad un IRC *server* si sceglie un canale tematico di discussione (su [www.efnet.net](http://www.efnet.net) ce ne sono più di 12000!).

Le conversazioni possono essere pubbliche o private. IRC non è un gioco! Le persone che si incontrano vanno trattate con cortesia, come se si dialogasse di persona o per telefono.

Nota: All'indirizzo:

[www.irchelp.org/irchelp/communication-research](http://www.irchelp.org/irchelp/communication-research)

sono disponibili diversi articoli sulla ricerca comunicazionale su *chat*.

Ogni *server* contiene alcuni canali etichettati in modo che si possa dedurre l'argomento di discussione.

Nota: su *Web* esistono siti che simulano IRC mediante *WebChat*, un programma *server* che permette il dialogo mediante pagine html create dinamicamente.

## Instant messaging

IRC può essere usato o per una conferenza a più voci, o, usando un canale riservato, per una conversazione privata.

Ma come fare a parlare con qualcuno, se non è connesso ad un canale?

Un servizio di *Instant Messaging* (Esempi: AOL *Instant Messenger* [www.aim.com](http://www.aim.com), ICQ [web.icq.com](http://web.icq.com), *MSN Messenger* [messenger.msn.com](http://messenger.msn.com)) permette di informare la rete che ci si è connessi, e disponibili a conversazioni.

In tutti i casi il servizio si basa su un *server* che registra una persona che si connette; ad ogni utente è assegnato un numero personale univoco che lo identifica in rete.

È possibile istruire il *server* su chi sono i nostri corrispondenti abituali, in modo che ci informi non appena si collegano.

ICQ (<http://web.icq.com>) è probabilmente il servizio di questo tipo più popolare: al dicembre 2001 ha circa 120.000.000 di iscritti.

## Newsgroup

I *newsgroup*, o bacheche elettroniche, sono raccolte di messaggi in cui si discute qualche argomento.

Esistono *newsgroup* dedicati genericamente alle scienze, alla politica, agli animali, o specificatamente alla cioccolata o agli scacchi giocati col *computer*.

Ogni *newsgroup* ospita di solito simultaneamente più filoni di discussione: ogni filone si chiama *thread*.

Per seguire un *newsgroup* ci sono diversi metodi:

- usare un cliente di *news* abilitato a ricevere un *newsfeed* da un *News server*; esistono clienti specifici, oppure nel caso delle *news* di **Internet** (*USENET*) si può usare un *browser Web* basato sul protocollo NNTP;
- usare un *server gateway*, che riceve un *newsfeed* e lo rilancia in forma di digesto agli iscritti ad una *mailing list*;
- usare siti specializzati nell'archiviare i *newsgroup*.

## Alcuni siti specializzati:

[www.deja.com](http://www.deja.com)

[groups.google.com](http://groups.google.com)

[www.mailgate.org](http://www.mailgate.org)

[www.mailgate.it](http://www.mailgate.it)

[www.newzbot.com](http://www.newzbot.com)

## USENET News

Una delle applicazioni più popolari di **Internet** è il sistema dei notiziari **USENET** che usa il protocollo NNTP; il numero di gruppi è alto (decine di migliaia) ed organizzato in gerarchie di notiziari:

- *Comp*: *Computer*, ricerca e industria informatica.
- *Sci*: Scienze fisiche ed ingegneristiche.
- *Humanities*: Letteratura e studi umanistici.
- *News*: Discussione su **USENET**.
- *Rec*: Attività ricreative, compresi sport e musica.
- *Misc*: Quel non appartiene a qualche altro gruppo.
- *Soc*: Socializzazione e argomenti sociali.
- *Talk*: Polemiche, dibattiti e discussioni.
- *Alt*: Gerarchia alternativa (ufficiosa).

Ogni gerarchia è divisa in sottoargomenti.

Esempio: *rec.sport* si occupa di sport, *rec.sport.basketball* di pallacanestro, e *rec.sport.basketball.women* di pallacanestro femminile.

- *Comp.ai*: Intelligenza artificiale.
- *Comp.os.ms-windows.video*: *Hardware e software* video per *Windows*.
- *Sci.geo.earthquakes*: Geologia, sismologia e terremoti.
- *News.lists*: Elenchi di gruppi **USENET**.
- *Rec.arts.poems*: Poesia.
- *Rec.food.chocolate*: Cioccolato.
- *Rec.humor.funny*: Barzellette.
- *Rec.music.folk*: Musica folk.
- *Rec.games.chess*: Scacchi.
- *Misc.jobs.offered*: Offerte di lavoro.
- *Soc.culture.italy*: Vita e cultura in Italia.
- *Soc.singles*: Persone sole e loro interessi.
- *Talk.rumors*: Pettegolezzi.
- *Alt.alien.visitors*: Incursioni di UFO.
- *Alt.tv.simpsons*: Bart e gli altri.
- *It.libri*: Recensioni di libri.
- *It.scacchi*: Scacchi in Italia.

## Netiquette per USENET

- Usare la firma automatica, ma non più lunga di 4 linee.
- Nei *newsgroup* ufficiali **USENET** (ovvero quelli diversi da *.ALT*) è considerato scorretto usare gli *attachments*.
- Non usare le *news* in alternativa alla *email* per raggiungere con un messaggio personale qualcuno che non risponde.
- Non pubblicare sulle *news* i messaggi di *email* di qualcun altro.
- Non pubblicare MAI messaggi di test sui *newsgroup* normali: per questo

scopo esitono i *newsgroup* di test come per esempio *alt.test*, *gnu.gnusenet.test*, *misc.test*.

- Usare la funzione di *Quoting* quando si risponde ad un articolo di qualcun altro:
  - In article <1232@foo.bar.com>, sharon@foo.bar.com wrote:  
> I agree, I think that basketweaving's really catching on,  
> particularly in Pennsylvania. Here's a list of every person  
> in PA that currently engages in it publicly:  
@centerline ... @rm{etc} ...
- Non abusare del *crossposting*.

## Suggerimenti di comportamento per mailing list e news

- Seguire regolarmente una *mailing list* o *newsgroup* per uno-due mesi prima d'inserire testi. Ciò permette di comprendere la cultura del gruppo.
- Considerare che ogni testo verrà letto da un ampio pubblico, incluso forse l'attuale o futuro datore di lavoro. Ricordarsi che sia le *mailing list* che i *newsgroup* vengono archiviati di frequente, e che i messaggi si conservano per molto tempo in un'area accessibile da molte persone.
- Assumere che ogni individuo esprima opinioni personali, senza impegnare la sua organizzazione (a meno che non venga esplicitamente detto il contrario).
- Messaggi ed articoli dovrebbero essere brevi e centrati sull'argomento. Non andare fuori tema, non ripetersi e non inviare messaggi né inserire estratti in un *posting* solo per sottolineare gli errori di battitura o i refusi degli altri.
- Se si scopre che un vostro messaggio personale è stato erroneamente recapitato ad una lista o gruppo, inviate un messaggio di scuse.
- Nel caso ci si trovi in disaccordo con una persona, meglio proseguire la discussione attraverso messaggi personali piuttosto che nella lista o gruppo. Se la discussione verte su un punto d'interesse generale, è sempre possibile riassumerla in seguito per gli altri.
- Mai farsi coinvolgere in una *flame war*: non rispondere a materiali incendiari.

## Come nasce un newsgroup

Un nuovo *newsgroup* nasce sempre da una votazione; le regole per la gerarchia italiana .it si trovano in:

[www.news.nic.it/news-it/status.html](http://www.news.nic.it/news-it/status.html)

- La richiesta di discussione (RFD) per la creazione di un gruppo va spedita all'indirizzo [rfd@news.nic.it](mailto:rfd@news.nic.it).
- La richiesta deve identificare proponente, nome e scopo del gruppo, e se è moderato.
- La RFD viene pubblicata su [it.news.annunci](http://it.news.annunci), ed altri gruppi rilevanti.
- Non vi è certezza che una RFD venga accettata. Le proposte non accettate

sono in [www.news.nic.it/news-it/status.html](http://www.news.nic.it/news-it/status.html).

- Il voto: passati almeno 15 giorni dalla pubblicazione di RFD, se il dibattito ha evidenziato interesse si inizia la procedura di voto che deve durare non più di 30 giorni.
- Per votare è necessario rispondere via *news* all'articolo contenente la procedura di voto in *it.news*
- Il gruppo verrà creato se i voti a favore sono pari o superiori a 100.

## Come funzionano le news USENET

Alcuni gruppi di interesse sono realizzati come *mailing list*. Per pubblicare un articolo lo si manda all'indirizzo della lista, che invia delle copie ad ogni indirizzo in lista.

*USENET* non viene realizzata con le *mailing list*. Ogni sito memorizza i messaggi in arrivo in un dato *directory*, per esempio *news*, con *sottodirectory* per comp, sci, eccetera che hanno *sottodirectory* come *news/comp/lang/java/misc*.

I programmi cliente lettori di *news* recuperano gli articoli a richiesta. Quindi ogni sito ha una sola copia di ogni articolo, indipendentemente dal numero di persone interessate. Dopo un po' gli articoli scadono e vengono rimossi dal disco.

Per aderire a *USENET*, un sito deve ottenere un *newsfeed* (alimentatore di *news*) da un altro sito su *USENET*. Si noti che essere su *Internet* non è né necessario né sufficiente per essere in *USENET*.

Periodicamente, ogni sito che vuole le *news* interroga i propri *newsfeed* per sapere se sono arrivate nuove notizie. Se sì, tali notizie vengono richieste, ricevute e memorizzate.

Può anche essere il *newsfeed*, invece del ricevente, a prendere l'iniziativa e stabilire il contatto quando ci sono abbastanza articoli nuovi.

Anni fa molti siti interrogavano i loro *newsfeed*, ma adesso la maggior parte usa l'altro modo.

I siti di solito non ricevono tutti i gruppi, perché i nuovi articoli superano i 2 GB al giorno, il che richiede una enorme quantità di spazio disco e di tempo di trasmissione.

## Web come servizio globale

Anche se non bisogna confondere *Internet* e *Web*, quest'ultimo è diventato negli ultimi anni il servizio principale che sussume e integra quasi tutti gli altri servizi *Internet*.

**Posta elettronica** su *Web*:

- si può usare un *browser* per leggere la propria *mbox*;
- si può attivare una *mbox* su un sito *Web* (esempio: *hotmail*).

*FTP/news/chat* su *Web*:

- si possono accedere siti ftp usando il *browser*;
- si può usare un *browser* per accedere un *newsfeed*;
- si può usare un *browser* per accedere un archivio di *news*;

- si può usare un *browser* per accedere una *chat-line*.

Radio/TV su *Web*:

- si può usare un *browser* per accedere canali radio/TV;
- si può usare un *browser* per accedere una WebCam.

Telefono o fax su *Web*:

- si può usare un *browser* per telefonare o mandare fax.

Commercio elettronico:

- si possono accedere siti per *shopping* o *home-banking*.

## Motori di ricerca

Cercare un documento, o più in generale una qualsiasi risorsa, sul *Web* senza conoscerne l'url è come cercare un ago in un pagliaio (anzi peggio viste le dimensioni del *Web*).

Per svolgere ricerche sul *Web* ci vengono in aiuto alcuni siti espressamente progettati e realizzati per tale scopo: i cosiddetti motori di ricerca.

Un motore di ricerca, o *spider*, è un sito con grandi capacità di connessione, immagazzinamento e calcolo che aiuta il navigatore *Web* a trovare l'url di siti che contengano le informazioni di cui ha bisogno.

Il motore di ricerca si basa su diversi componenti:

- un programma che periodicamente legge e cataloga le pagine di **Internet**.
- Un *database* delle pagine catalogate.
- Una interfaccia di interrogazione.

Quando un utente (navigatore *Web*) si collega ad un motore di ricerca lo interroga per trovare tutti i documenti (e relativi url) che contengono alcune parole chiave da lui stesso fornite tramite l'interfaccia di interrogazione.

Ai motori di ricerca sarà dedicato uno dei moduli di approfondimento.

## Esempi di motori di ricerca:

Motori di ricerca internazionali:

- [www.google.com](http://www.google.com)
- [www.altavista.com](http://www.altavista.com)
- [www.yahoo.com](http://www.yahoo.com)
- [www.excite.com](http://www.excite.com)

Motori di ricerca italiani:

- [www.arianna.it](http://www.arianna.it)
- [www.virgilio.it](http://www.virgilio.it)

Alcuni riferimenti utili:

- [www.searchenginewatch.com](http://www.searchenginewatch.com)
- [www.webposition.com](http://www.webposition.com)

- [www.cisi.unito.it/pointes/motori](http://www.cisi.unito.it/pointes/motori)



# Installazione e configurazione di applicazioni

Luciano Margara

## Installare e configurare un browser

Come già spiegato nelle precedenti parti introduttive di questo modulo un *browser* è una applicazione (lato *client*) che ha lo scopo principale di permettere all'utente di navigare in *Internet*. Spesso l'applicazione *browser* fornisce alcune funzionalità aggiuntive, come ad esempio la gestione della posta elettronica.

In questo approfondimento descriveremo in dettaglio le procedure di installazione e di configurazione di un *browser*. Per far ciò utilizzeremo il *browser Internet Explorer* (IE - versione 6 - in italiano) di *Microsoft*.

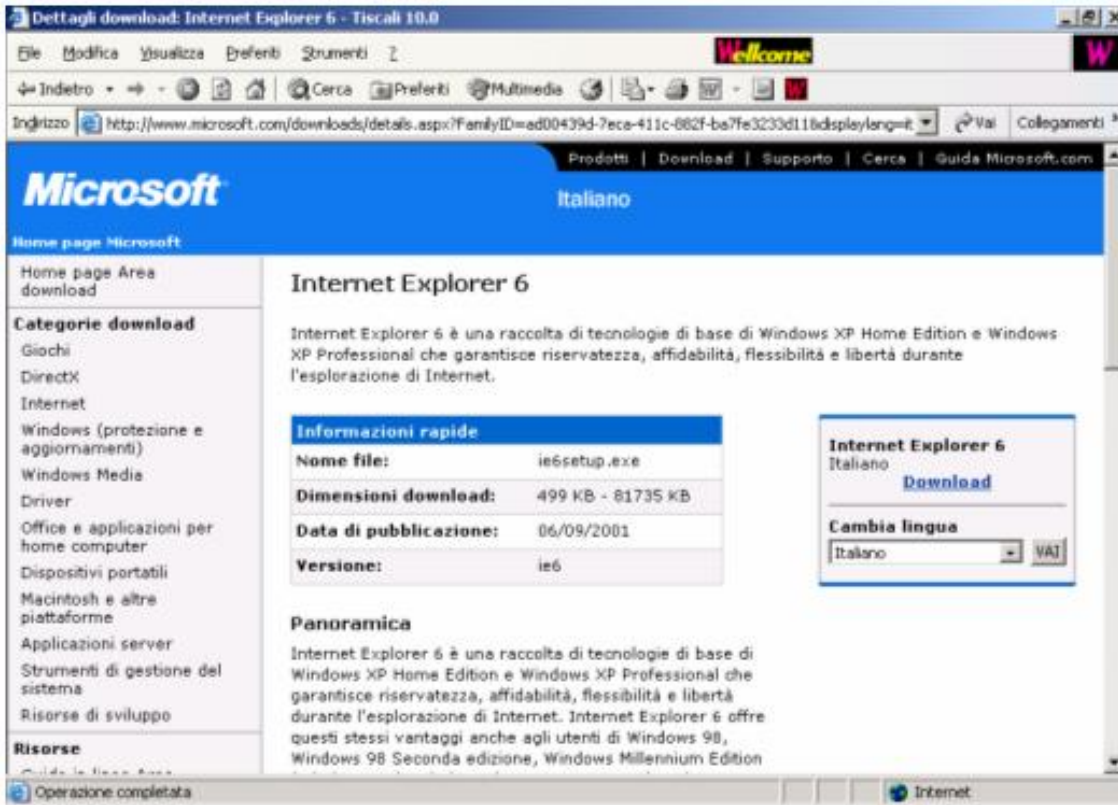
Questa scelta è stata suggerita dal fatto che la stragrande maggioranza dei navigatori usa effettivamente IE come *browser* predefinito.

Ci sentiamo però di affermare che sul mercato esistano altri *browser* equivalenti a IE (esempio: *Netscape Navigator*) che, sul piano della funzionalità e della usabilità, non hanno nulla da invidiare a IE.

Lo studio di IE semplicemente ci permetterà di soddisfare le aspettative di un numero maggiore di persone.

## Installazione di IE

IE è gratis. Può essere scaricato direttamente da *Internet* (ad esempio dal sito *Internet* di *Microsoft*. Figura 1.)



Sito della Microsoft da cui scaricare IE6

Figura 1. Sito della *Microsoft* da cui scaricare IE6.

o essere ad esempio contenuto in CD dato in omaggio con l'acquisto di una rivista informatica specializzata.

Prima di procedere con l'installazione conviene:

controllare che il proprio calcolatore abbia i requisiti *hardware* e *software* minimi consigliati. IE non necessita di grandi quantità di RAM, di processori estremamente potenti o di un disco particolarmente capiente. Solitamente già il fatto che il vostro calcolatore supporti adeguatamente il vostro sistema operativo dovrebbe permettervi di utilizzare IE senza problemi.

Disporre di un *software* antivirus installato sul proprio calcolatore (non si sa mai durante la navigazione che incontri si possono fare).

Avviando il programma di installazione di IE (che generalmente ha un nome tipo ie6setup.exe) all'utente viene chiesto di accettare la licenza di utilizzo del *software* (Figura 2).



Finestra di accettazione della licenza di Internet Explorer

Figura 2. Accettazione della licenza d'uso.

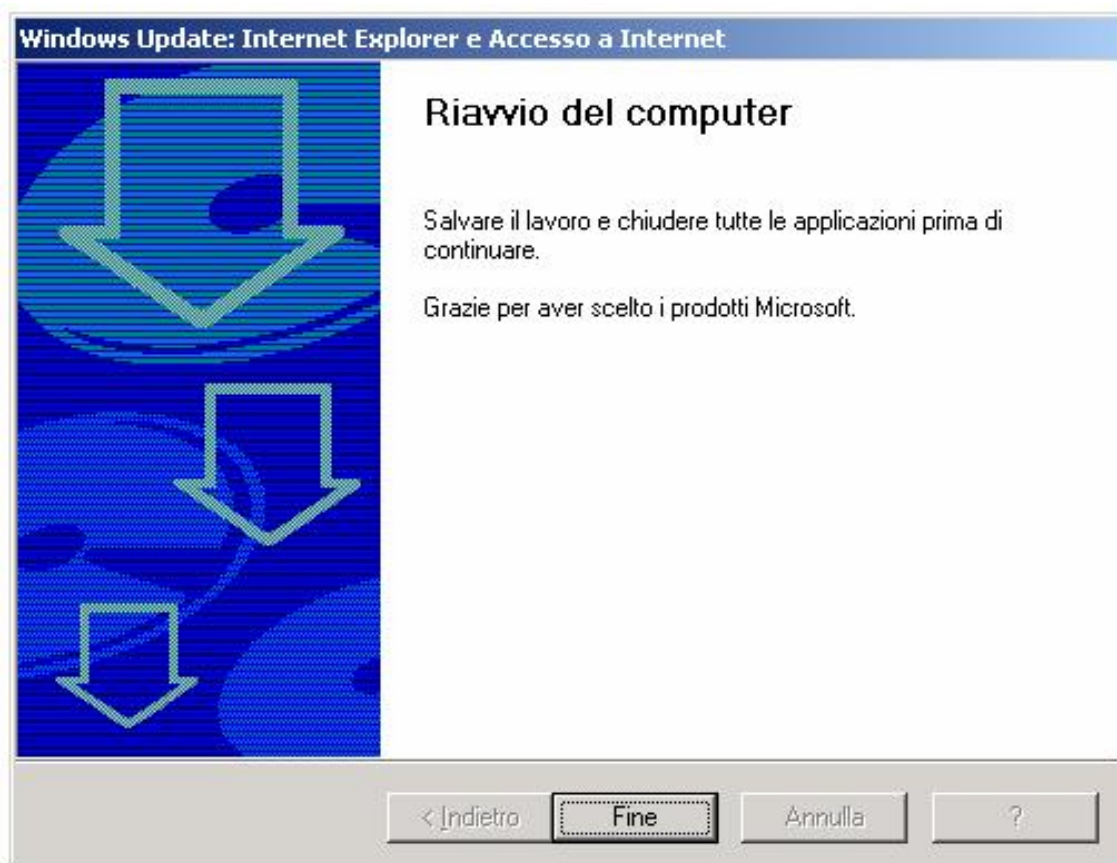
A questo punto l'installazione procede in modo leggermente diverso a seconda del sistema operativo posseduto e delle versioni di IE eventualmente già installate (se si possiede una versione di IE precedente alla versione 6 si procederà ad un aggiornamento invece che ad una installazione completa). Con *Windows 2000* o *XP* non c'è altro da fare. L'installazione termina senza dover compiere nessun'altra scelta (Figura 3).



Finestra di fine installazione

Figura 3. Fine installazione.

Cliccando su Avanti otterremo una finestra che ci richiede di avviare il *computer* (Figura 4).



Finestra di riavvio del computer

Figura 4. Riavvio del *computer*.

Con versioni precedenti di sistema operativo occorre scegliere tra diversi tipi di installazione.

Vediamo cosa succede ad esempio con *Windows 98*.

All'utente viene chiesto di scegliere tra una installazione tipica (che risponde alle esigenze *standard* dell'utente medio) e una pagina iniziale delle singole componenti di IE.

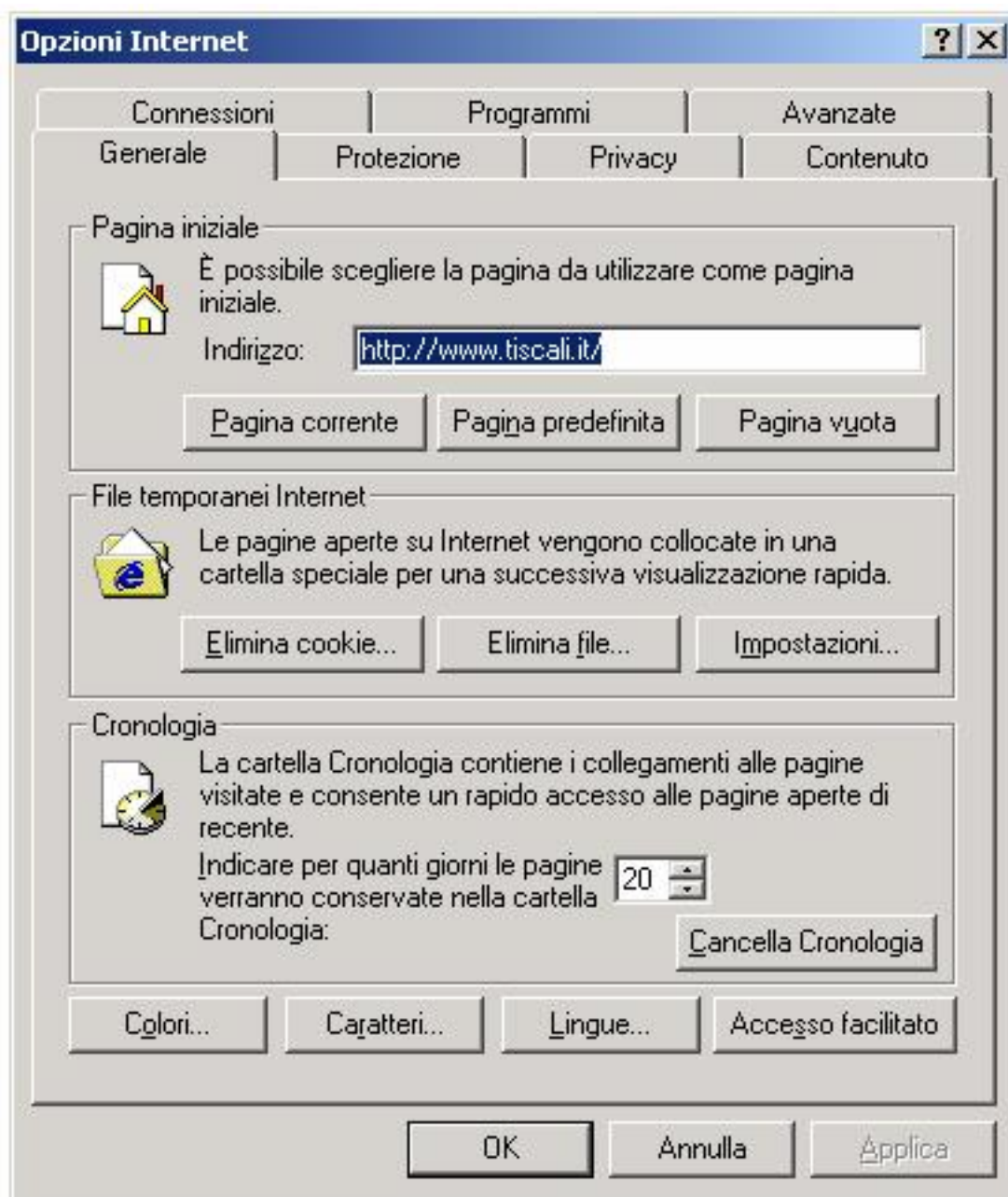
L'installazione tipica include: il *browser* vero e proprio, *Outlook express* (applicazione che gestisce la posta elettronica), *Windows Media Player* e altri componenti multimediali avanzati.

La pagina iniziale permette di selezionare a piacimento quali moduli di IE installare (ad esempio se già abbiamo un programma di posta elettronica possiamo decidere di non installare *Outlook Express*) e in che cartella del nostro *file system* memorizzarli. Ci sono tre installazioni preconfezionate da cui partire per poi raffinare la nostra scelta selezionando o deselezionando a piacimento i singoli componenti: minima, tipica e completa.

Nel seguito di questo approfondimento assumeremo di aver scelto e correttamente portato a termine una installazione tipica.

## Opzioni Internet di IE

Per configurare IE in modo da ottenere un *browser* adatto alle proprie esigenze occorre selezionare la voce *Opzioni Internet* nel menu Strumenti. A questo punto ci troveremo davanti ad un pannello che ci permetterà di operare le nostre scelte. Il pannello (Figura 5) contiene al suo interno 7 cartelle (Connessioni, Programmi, Avanzate, Generale, Protezione, *Privacy* e Contenuto) ognuna delle quali fornisce all'utente una serie di opzioni che guideranno IE durante la navigazione. Vediamole in dettaglio.



Pannello per configurare le caratteristiche principali di Internet Explorer



Figura 5. Strumenti > Opzioni *Internet* > Generale

## Opzioni Internet > Generale

Nel pannello denominato Generale (Figura 5) possiamo configurare le voci seguenti:

**Pagina iniziale.** La pagina iniziale (specificata fornendo la sua url) è la pagina *Internet* (scelta da noi) che verrà scaricata (se non già presente in memoria *cache*) e visualizzata ogni qual volta attiveremo IE sul nostro calcolatore oppure cliccheremo sulla barra dei pulsanti (*toolbar*) in corrispondenza dell'icona a forma di casetta (*home page*). Scelte tipiche per la pagina iniziale sono:

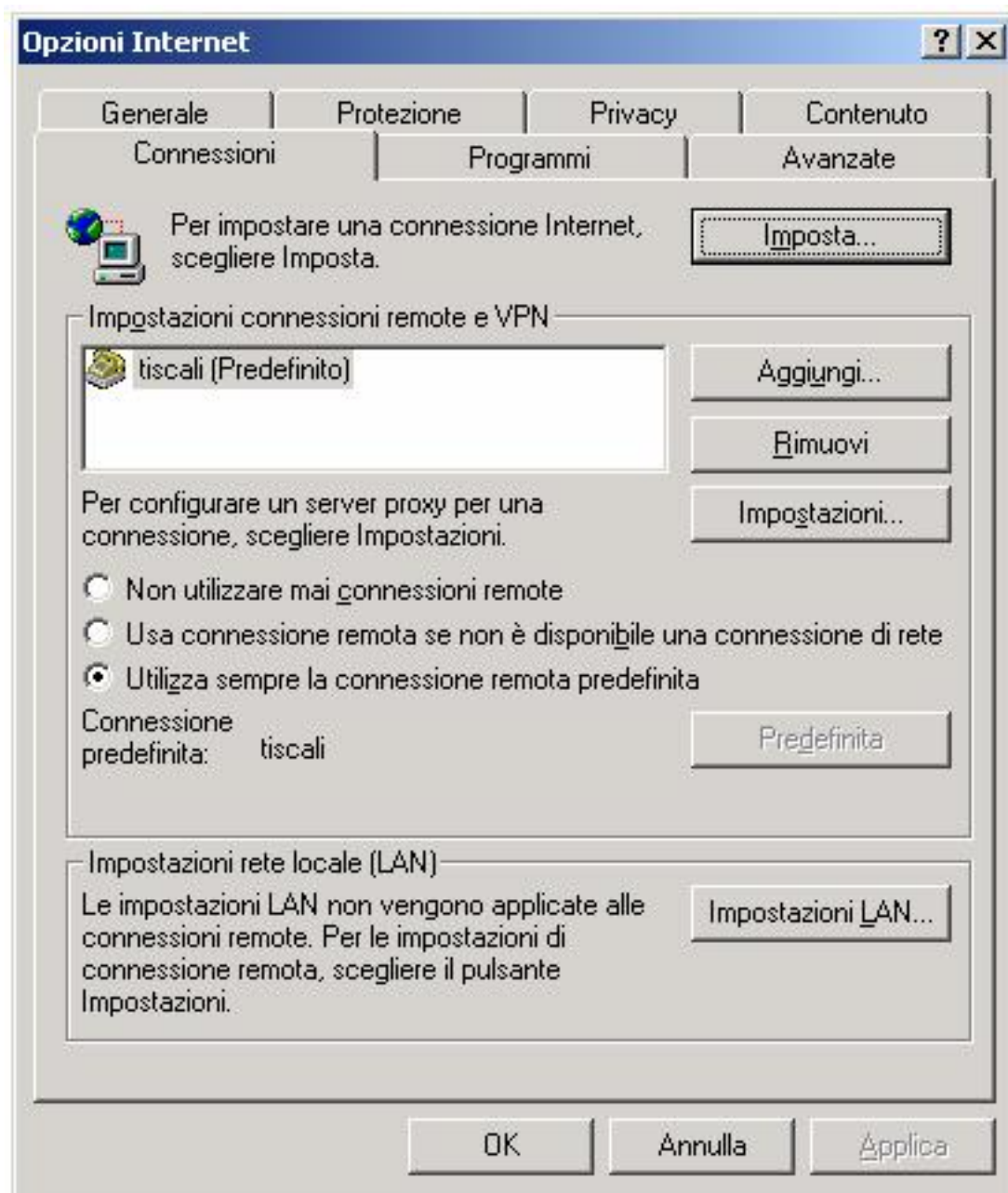
- il proprio sito *Internet*;
- un motore di ricerca;
- il sito della organizzazione di cui facciamo parte;
- un portale.

File Temporanei. IE (ma anche gli altri *browser*) memorizzano alcuni dati (principalmente *cookie* e pagine recentemente visitate) in una zona di memoria detta memoria *cache*. Questa operazione permette di utilizzare tali dati in futuro senza doverli scaricare nuovamente dalla rete.

È possibile eliminare volontariamente i *cookie* e i *file* precedentemente memorizzati cliccando sui relativi pulsanti. Quando la zona di memoria dedicata ai *file* temporanei è esaurita (piena) il *browser* provvederà automaticamente a eliminare le informazioni più vecchie in essa contenute.

Cronologia. IE mantiene memorizzate le url delle pagine visitate di recente. In questo modo è possibile selezionare una pagina da questo elenco in modo veloce senza doversi ricordare l'url a memoria. È possibile dire al *browser* per quanti giorni tenere memorizzata in questo elenco l'url di una pagina visitata.

Nota: IE non memorizza il contenuto delle pagine, ma solamente la loro url. Se ad esempio vogliamo navigare su un sito di cui non ci ricordiamo la url, ma che sappiamo aver visitato alcuni giorni prima, possiamo risalire alla sua url ricercandola tra le url dei siti visitati di recente.



pannello per configurare le connessioni a Internet

Figura 6. Strumenti > Opzioni *Internet* > Connessioni

## Opzioni Internet > Connessioni

Nello spazio denominato Impostazioni connessioni remote e VPN viene mostrato l'elenco delle connessioni a *Internet* che l'utente si è creato in precedenza. Ogni connessione permette l'accesso remoto a *Internet* (il collegamento vero e proprio) con uno specifico *Internet Service Provider*.

Usando i pulsanti Aggiungi, Rimuovi e Imposta è possibile gestire le connessioni



remote (aggiungerne una nuova oppure modificarne/cancellarne una già esistente). In Figura 7 è visualizzato il pannello relativo a una connessione remota.



pannello per configurare una connessione remota

Figura 7. Esempio di connessione remota.

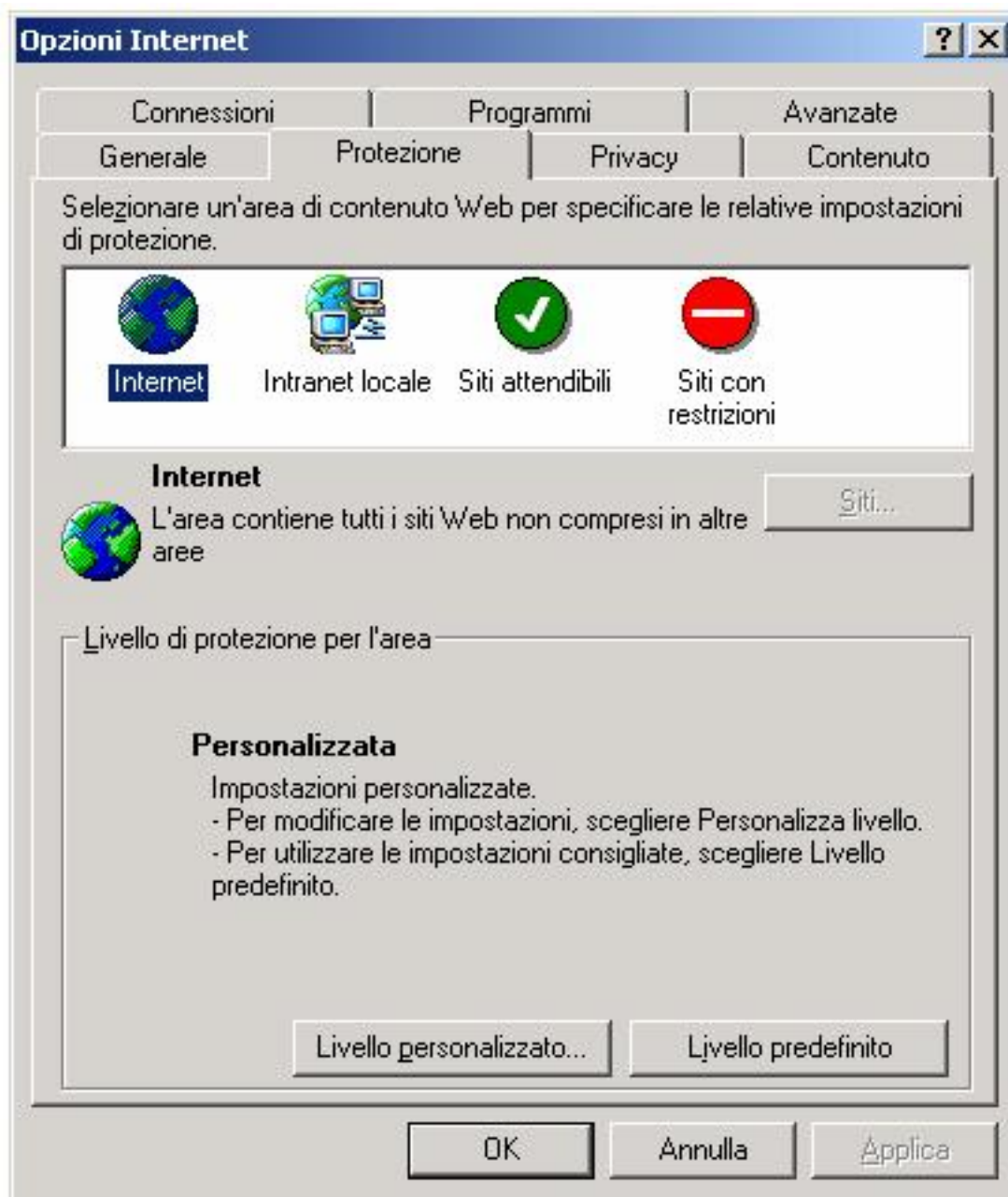
L'utente può decidere inoltre se non attivare mai la connessione remota in automatico, se attivarla quando non è disponibile un accesso alla rete locale, oppure se utilizzare sempre (in automatico) la connessione remota battezzata come predefinita.

L'utente infine può configurare alcuni parametri riguardanti il collegamento con la rete locale (se esiste) cliccando su Impostazioni LAN.

## Opzioni Internet > Protezione

Quando si naviga in *Internet* il nostro calcolatore e i dati in esso contenuti possono trovarsi in situazioni di pericolo.

Il pannello protezione suddivide i siti *Web* in 4 categorie e per ognuna di esse permette di specificare il livello di protezione che vogliamo garantire al nostro *computer* e ai dati in esso contenuti.



pannello per configurare le opzioni di protezione

Figura 8. Strumenti > Opzioni *Internet* > Protezione

I siti *Internet* sono suddivisi (Figura 8) in quattro categorie:

- Intranet locale: include tutti i siti *Web* compresi nella rete Intranet della società (se tale rete esiste).
- Siti attendibili: include i siti *Web* considerati non pericolosi per il *computer* o per i dati in esso contenuti.
- Siti con restrizioni: include siti *Web* potenzialmente dannosi per il *computer* o per i dati in esso contenuti.

- **Internet:** include i siti *Web* che non appartengono alle altre 3 categorie. Per ognuna delle 4 categorie sopra descritte l'utente può selezionare due modalità di protezione: personalizzata e predefinita.

Come al solito la modalità personalizzata dà la possibilità all'utente di decidere in modo dettagliato il livello di protezione del suo *computer* nei confronti dei siti che appartengono a ciascuna categoria.

La modalità predefinita invece consente di modificare il livello di protezione in modo molto semplice e intuitivo semplicemente agendo su un cursore grafico.

## Opzioni Internet > Privacy

La *privacy* del navigatore può essere messa in pericolo attraverso diversi meccanismi di interazione tra siti visitati e calcolatore dell'utente.

Il metodo forse più diffuso per violare la *privacy* o più semplicemente per raccogliere informazioni sul profilo del navigatore è l'utilizzo dei cosiddetti *cookie* (dall'inglese biscottino).

Un *cookie* è un'informazione che il sito su cui navighiamo cerca di memorizzare sul nostro *computer*. Questa informazione può essere di diversa natura e servire a molti scopi. Non necessariamente il *cookie* è dannoso per il navigatore. Ad esempio il *cookie* può essere utilizzato da un sito per memorizzare informazioni sui gusti del navigatore in modo tale da poter personalizzare la pagina *Web* che sarà visualizzata in una eventuale altra visita del sito. I siti che forniscono servizi di commercio elettronico ad esempio possono in questo modo ricordarsi quali merci ha già acquistato il cliente e proporre altre di tipo simile.



Pannello per configurare le opzioni di privacy

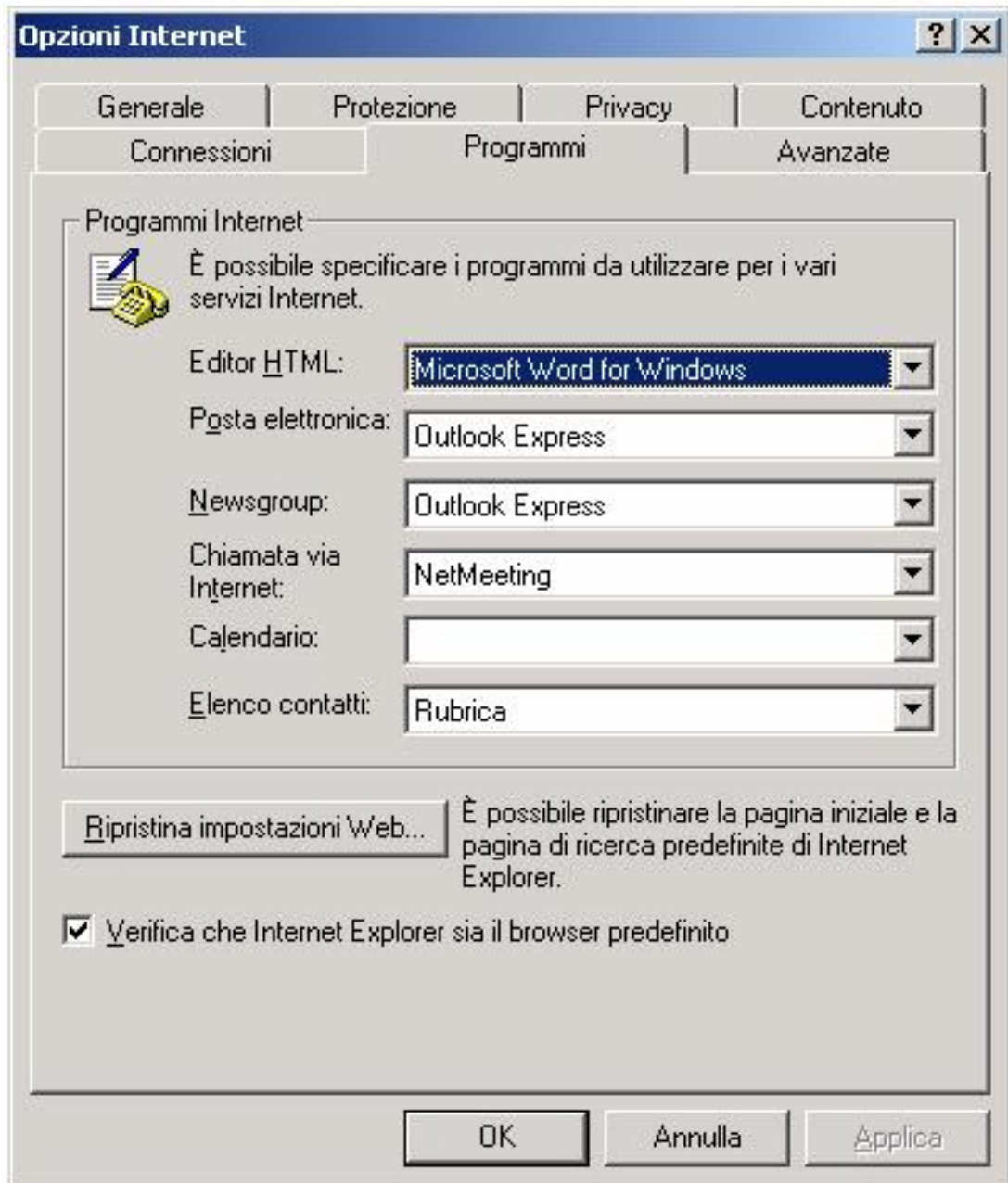
Figura 9. Strumenti > Opzioni *Internet* > *Privacy*

Il pannello *Privacy* (Figura 9) permette di inibire completamente la memorizzazione dei *cookie* sul nostro computer oppure di essere avvisati nel momento in cui si tenta di memorizzare un *cookie*. Anche in questo caso esistono configurazioni predefinite per aiutare l'utente non esperto.

## Opzioni Internet > Programmi

Utilizzando il pannello Programmi (Figura 10) l'utente può specificare un insieme di

applicazioni da usare in corrispondenza dei servizi di *Internet* richiesti.



pannello per scegliere i programmi da utilizzare per i vari servizi Internet

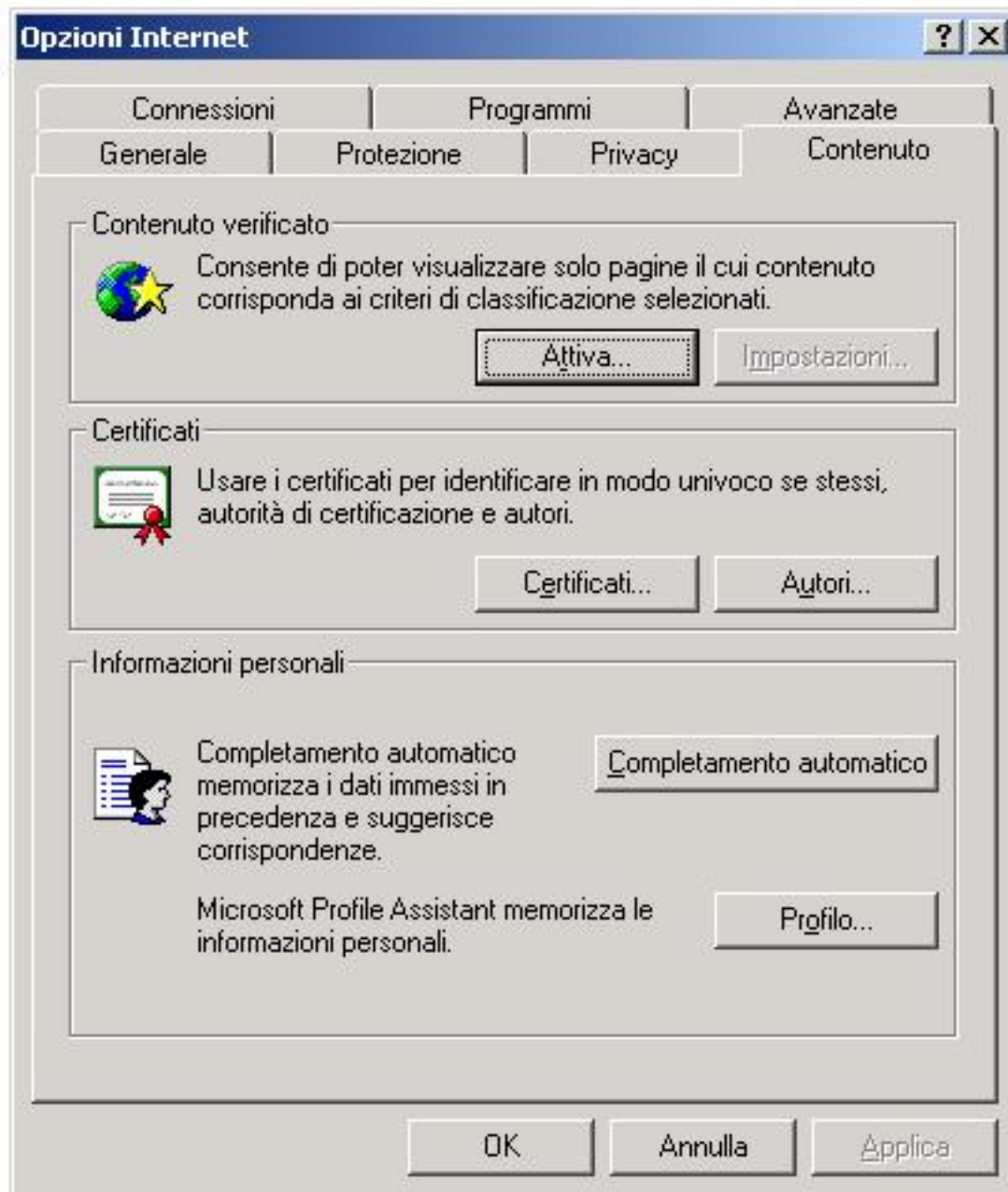
Figura 10. Strumenti > Opzioni *Internet* > Programmi

Esempio. Nel campo Posta elettronica possiamo specificare un cliente di posta elettronica da usare nel caso in cui durante la navigazione di un sito ci capiti di cliccare su un *link* ad un indirizzo di posta elettronica. Automaticamente IE attiverà il cliente di posta indicato nelle preferenze (in questo caso: *Outlook Express*).

## Opzioni Internet > Contenuto.

Utilizzando il pannello Opzioni *Internet* > Contenuto (Figura 11) l'utente può:

- decidere di impedire l'accesso/navigazione su alcuni siti particolari;
- utilizzare lo strumento dei certificati per identificare in modo univoco e sicuro alcuni soggetti sulla rete (incluso se stesso);
- autorizzare IE a riempire in modo automatico eventuali moduli incontrati durante la navigazione.



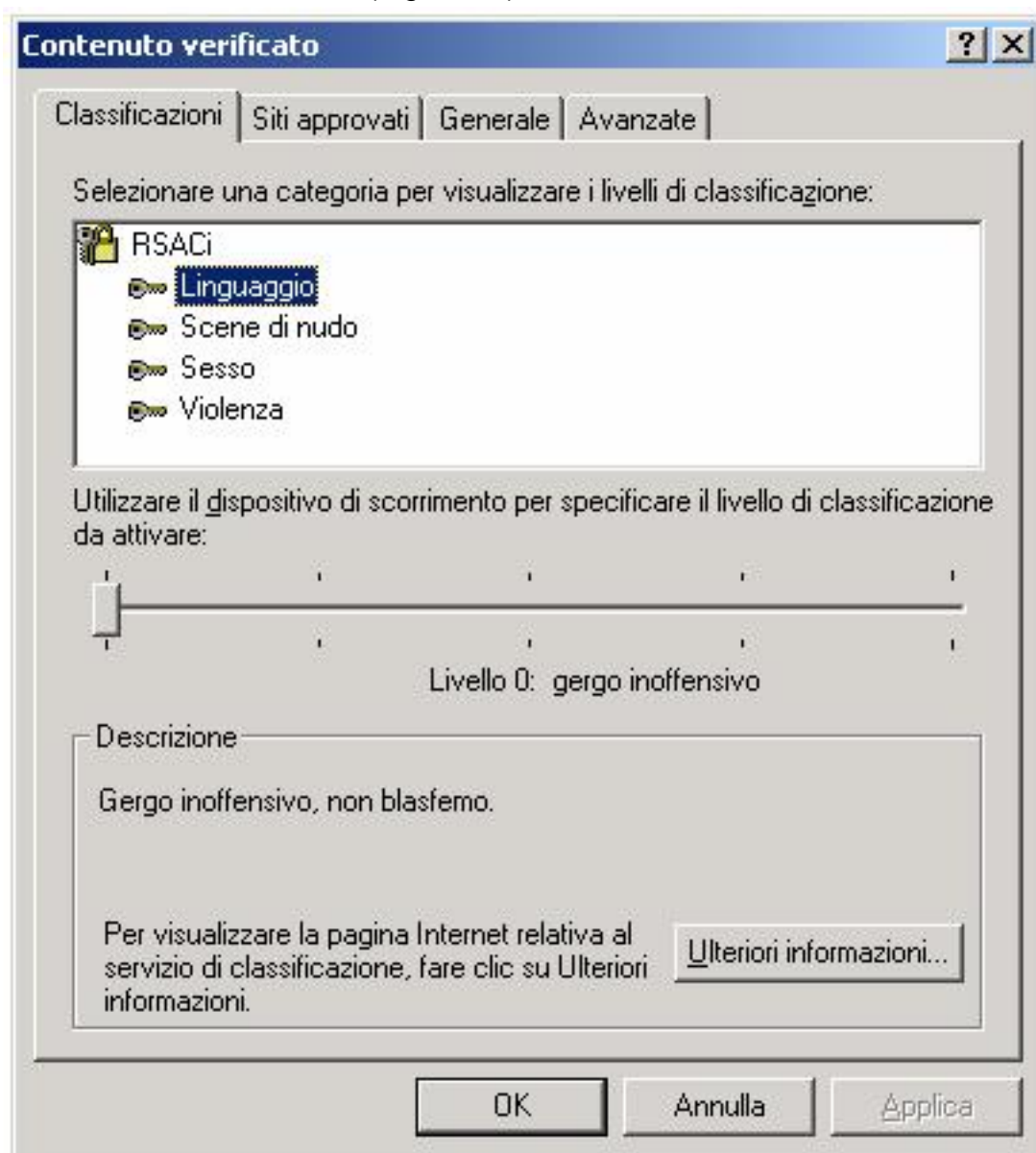
pannello per configurare le opzioni di limitazione alla navigazione

Figura 11. Strumenti > Opzioni *Internet* > Contenuto

Contenuto verificato. Consente di:

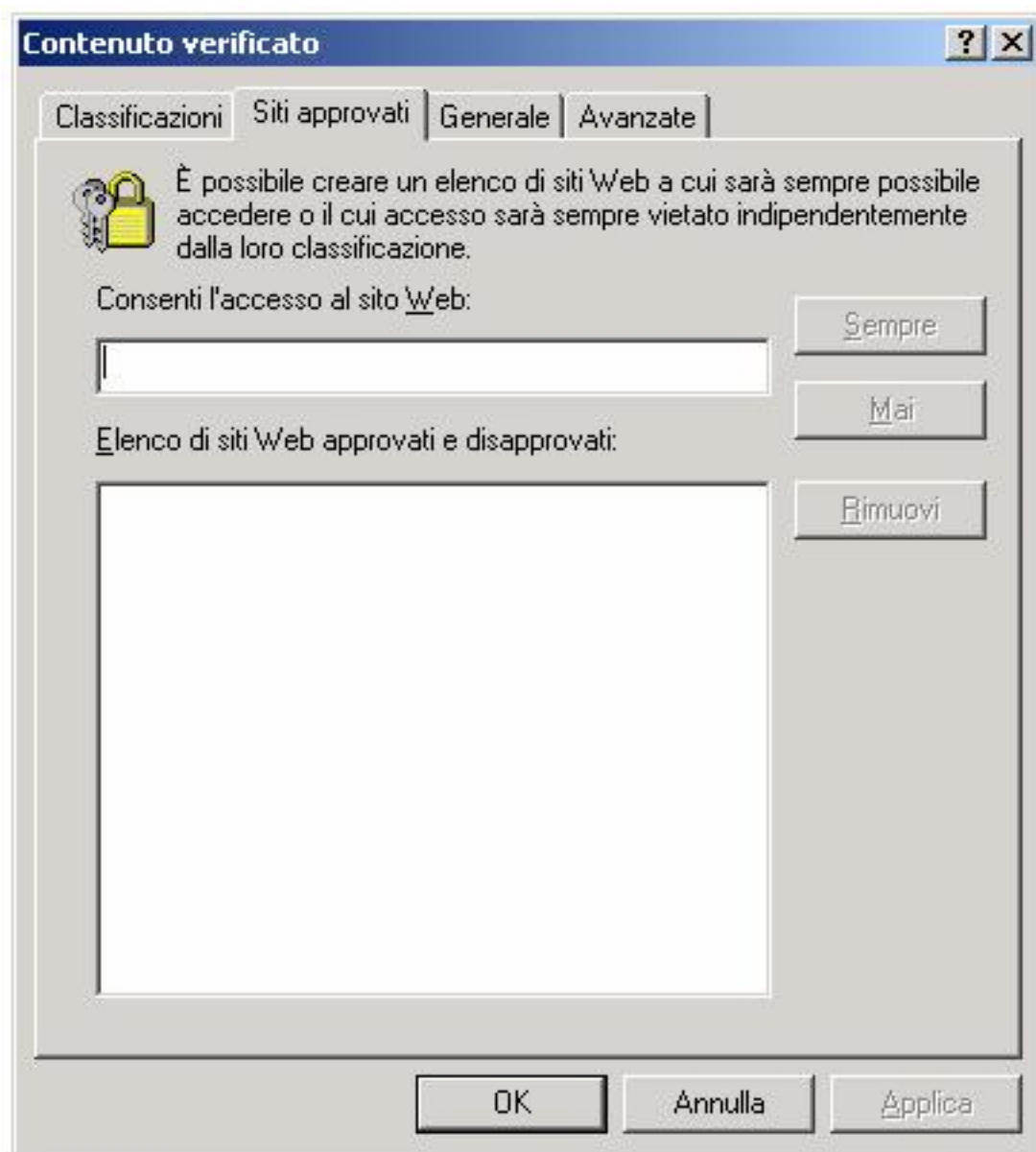


- impedire l'accesso a determinati siti, in base ad alcuni parametri quali la presenza di immagini pornografiche, di linguaggio volgare e così via, in base alla classificazione proposta dall'ente RSAC (Figura 12).
- Impedire sempre o consentire sempre l'accesso ad una lista di siti esplicitamente indicati dall'utente indipendentemente dalla classificazione RSAC (Figura 13).
- Visualizzare siti senza classificazione. Definire una *password* (password del supervisore) da inserire per consentire comunque la visualizzazione di siti altrimenti proibita. Cambiare il metodo di classificazione RSAC proposto come *default* da IE. Queste opzioni sono visualizzate in Figura 14.
- Gestire in modo avanzato (con l'ausilio di agenzie esterne) il sistema di classificazione dei siti (Figura 15).



pannello per configurare le opzioni di classificazione dei siti

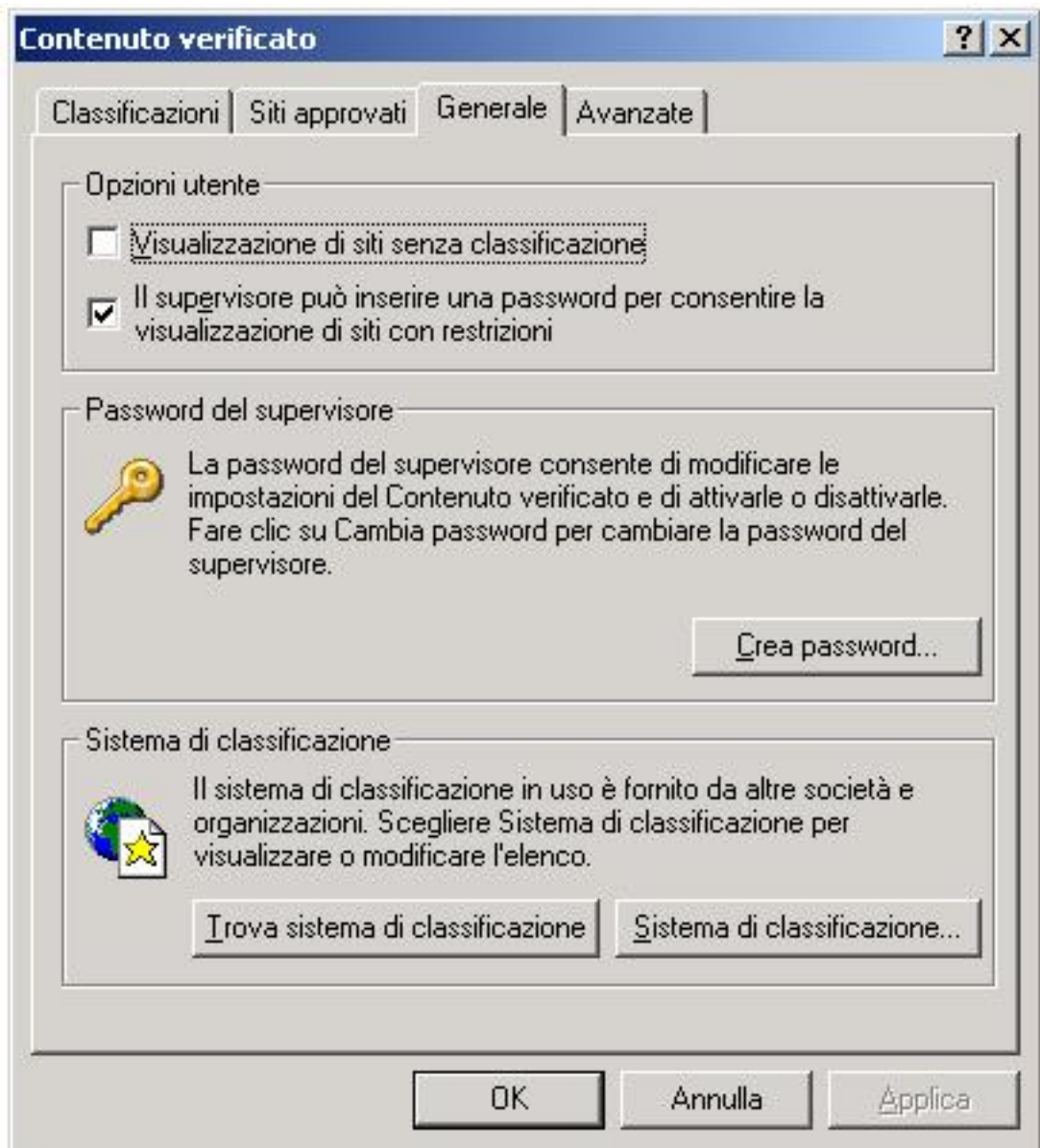
Figura 12. Strumenti > Opzioni *Internet* > Contenuto > Contenuto verificato > Classificazioni



elenco dei siti esplicitamente approvati o disapprovati dall'utente

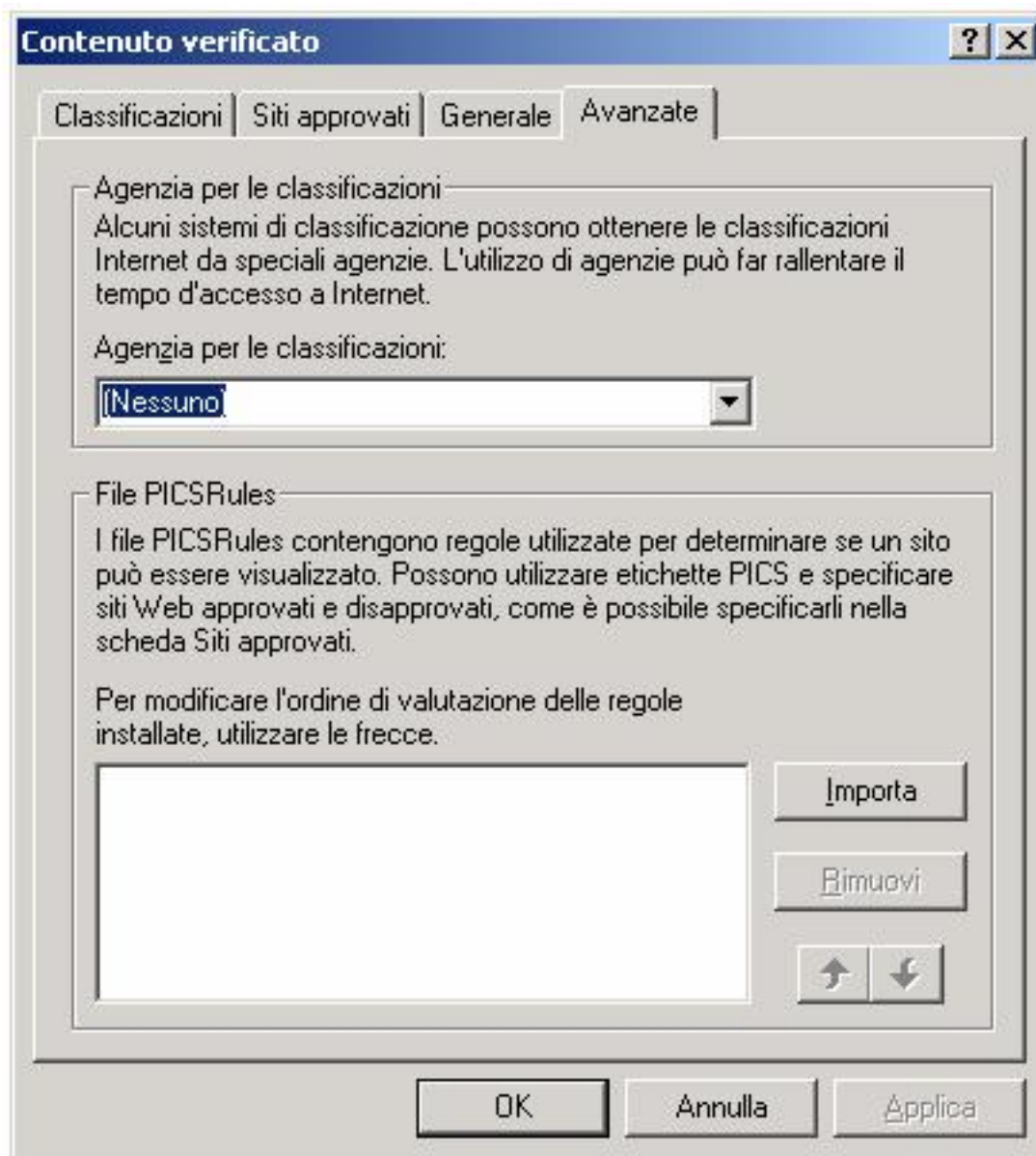
Figura 13. Strumenti > Opzioni *Internet* > Contenuto > Contenuto verificato > Siti approvati





pannello per configurare le opzioni generali circa la verifica del contenuto dei siti

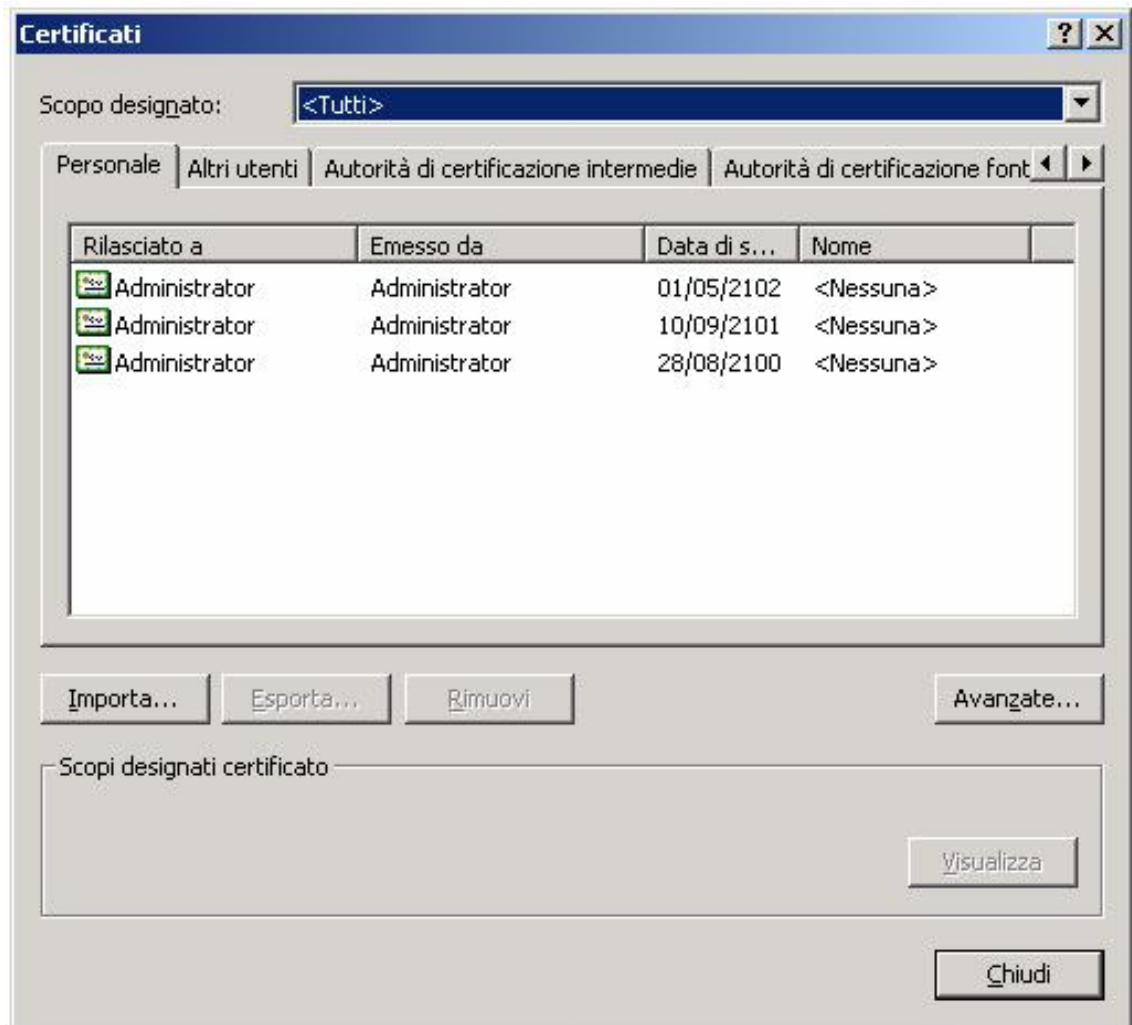
Figura 14. Strumenti > Opzioni *Internet* > Contenuto > Contenuto verificato > Generale



pannello per configurare le opzioni avanzate circa la verifica del contenuto dei siti

Figura 15. Strumenti > Opzioni *Internet* > Contenuto > Contenuto verificato > Avanzate

Certificati. Consente, attraverso un meccanismo di certificazione, di battezzare come sicuri vari soggetti ad esempio l'utente, altre persone, enti e autorità varie (Figura 16).



pannello per configurare le opzioni generali circa l'utilizzo dei certificati

Figura 16. Strumenti > Opzioni *Internet* > Contenuto > Certificati

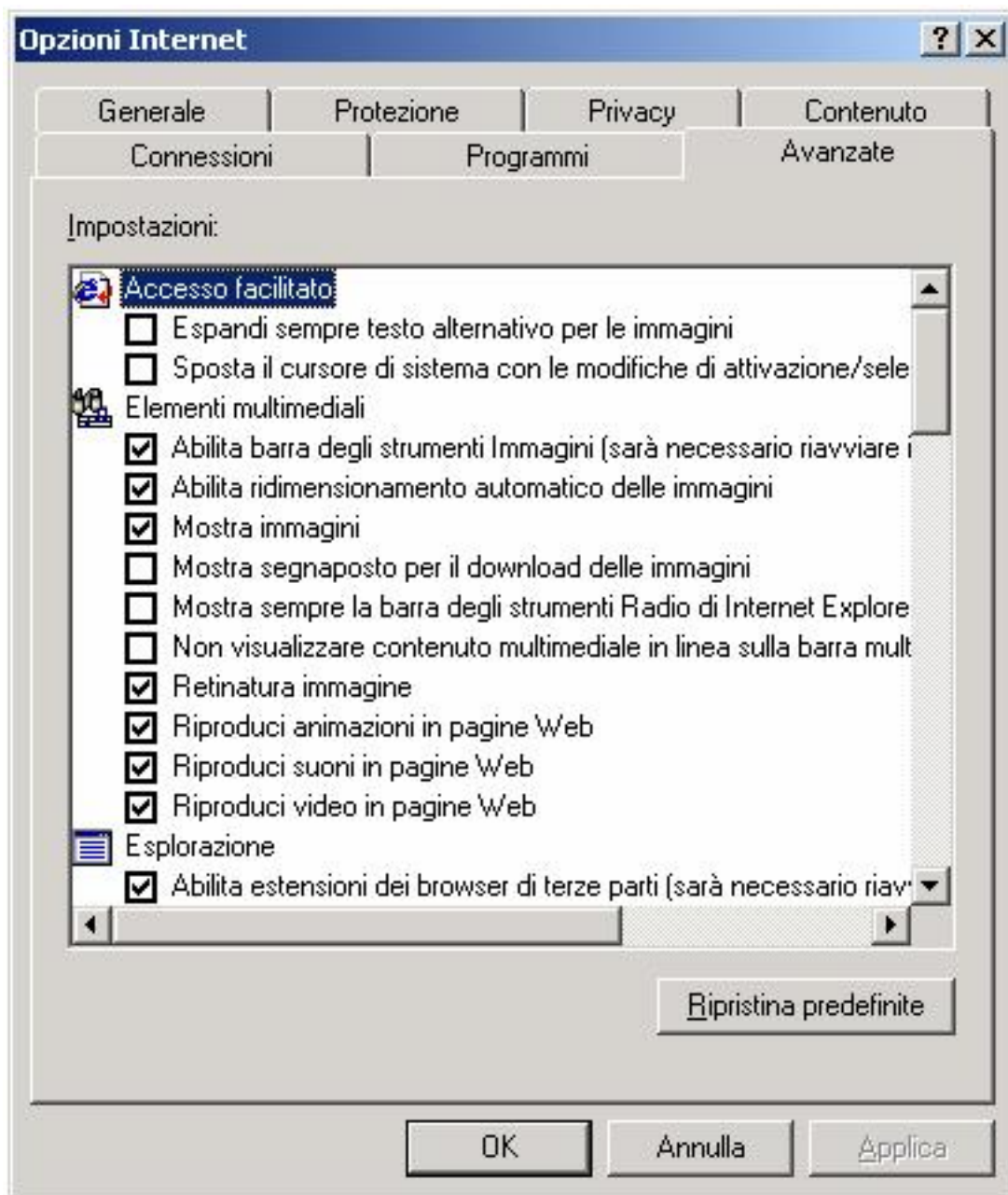
Informazioni personali. Consente di creare un profilo utente da utilizzare quando alcuni siti *Web* vi presentano dei moduli da completare mediante inserimento di dati personali. Una volta creato il proprio profilo è possibile dare istruzioni a IE di riempire automaticamente il modulo con i propri dati senza dover ogni volta inserirli a mano.



pannello per configurare le opzioni relative alle informazioni personali dell'utente

Figura 17. Strumenti > Opzioni *Internet* > Contenuto > Informazioni personali

## Opzioni Internet > Avanzate.



pannello per configurare le opzioni avanzate di Internet Explorer

Figura 18. Strumenti > Opzioni *Internet* > Avanzate.

Nel pannello Avanzate (Figura 18) possiamo configurare un gran numero di altre opzioni che regolano il comportamento di IE in occasioni particolari. La gestione delle opzioni avanzate va oltre lo scopo di questo corso. Per avere una spiegazione sintetica su ogni voce configurabile in questo pannello, cliccate sul punto interrogativo situato nell'angolo in alto a destra e poi cliccate sulla voce desiderata accedendo in questo modo all'*help on-line* fornito da IE.

# Approfondimento

## Motori di Ricerca

Luciano Margara

12.2.2 (Usare tool e utilità Internet quali: e-mail, browser, motori di ricerca, newsgroup, list server, chat, trasferimento di file)

### Introduzione

Come abbiamo detto nella parte introduttiva di questo modulo, il *Web* può essere visto come un immenso documento ipertestuale multimediale distribuito. Per accedere ad una pagina (o a una qualsiasi altra risorsa *Internet* (file multimediali, documenti in vari formati digitali, eccetera) pubblicata sul *Web* occorre conoscere il suo indirizzo (url). Nella maggior parte dei casi però l'utente *Web* non conosce l'url della risorsa *Internet* a cui vuole accedere. Solitamente il navigatore ha a disposizione soltanto informazioni generiche circa l'argomento della sua ricerca.

Alla fine dell'anno 2001 il *Web* conteneva circa 4 miliardi di documenti (pagine *Web*). In media, ogni giorno, vengono pubblicate sul *Web* 7 milioni di nuove pagine.

Immaginatevi di essere uno studente universitario che vuole utilizzare *Internet* per reperire informazioni sull'argomento trattato nella sua tesi di laurea, oppure immaginatevi di voler trovare l'offerta più conveniente per un mutuo prima casa, o ancora di voler acquistare un biglietto aereo per le vostre vacanze. Come è possibile scovare l'indirizzo delle pagine che contengono queste informazioni? Questo è in sintesi il problema che affronteremo in questo approfondimento.

### Cosa è un motore di ricerca

Risulta evidente dalle considerazioni appena fatte la necessità di disporre di uno strumento automatico che ci metta in grado di ricercare informazioni in *Internet* in modo veloce, accurato ed economico. Uno strumento che svolga la stessa funzione dell'elenco telefonico quando dobbiamo cercare un numero di telefono o di un catalogo di libri quando dobbiamo cercare un volume in una biblioteca. Nel caso di *Internet* la situazione è però molto più complessa in quanto:

- *Internet* contiene un numero gigantesco di documenti che si trovano fisicamente distribuiti su tutto il pianeta;
- ogni giorno moltissimi documenti cessano di esistere e moltissimi altri vengono pubblicati per la prima volta;
- le risorse in *Internet* sono di molti tipi e hanno una struttura disomogenea;
- non esiste una chiave sola di ricerca come accade nel caso dell'elenco telefonico.

Attualmente gli strumenti di ricerca che meglio soddisfano i requisiti di velocità accuratezza ed economicità sono i cosiddetti motori di ricerca.

In estrema sintesi un motore di ricerca è un sito *Internet* che, partendo dalla descrizione dell'argomento oggetto della ricerca, è in grado di restituire un elenco (possibilmente vuoto se la ricerca non va a buon fine) di pagine (e relativi url) che



hanno a che fare con quell'argomento.

- Come può un sito *Internet* ricercare informazioni sul *Web*?
- Come è possibile descrivere ciò che vogliamo ricercare?
- Come si fa a stabilire che una certa pagina *Web* è interessante per la nostra ricerca?

Queste sono solo alcune delle domande che dovrebbero nascere spontanee dalla definizione che abbiamo dato di motore di ricerca e alle quali cercheremo di rispondere in questo approfondimento.

## Motori di ricerca e directory

Gli strumenti di ricerca disponibili su *Internet* sono sostanzialmente di due tipi: **directory** e motori di ricerca veri e propri (a volte chiamati anche *spider*).

Spesso viene usato impropriamente il termine motore di ricerca anche per le *directory*. In realtà *directory* e motori di ricerca sono due concetti distinti. Nella tabella riportata qui di seguito elenchiamo alcune caratteristiche basilari delle *directory* e dei motori di ricerca.

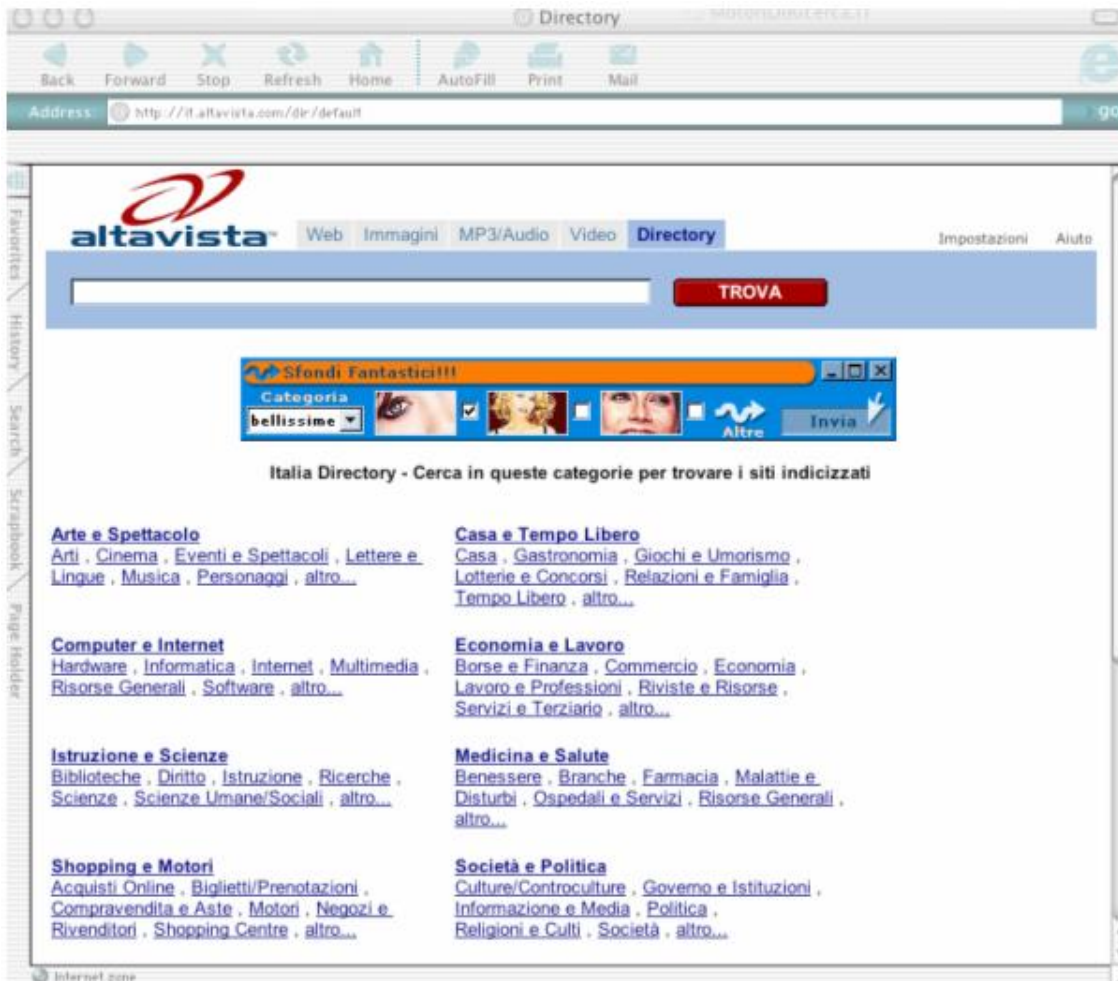
Motori veri e propri ( <i>spiders</i> )	Directory
Automatici	Risultato di lavoro umano
Indicizzano url	Classificano interi siti
Mirano a indicizzare tutto il Web	Mirano a diventare guide alle migliori risorse Web
Si basano su algoritmi matematici	Si basano su giudizi di valore qualitativi
Interrogabili attraverso combinazioni di parole chiave	Percorribili per categorie e sotto-categorie

La confusione tra *directory* e motori nasce dal fatto che quasi tutti i motori ospitano una *directory* (Altavista ospita *Looksmart*, Excite ospita *Excite's Guide* ...) e quasi tutte le *directory* sono interrogabili per navigazione ma anche per parole chiave (anche se meno efficacemente).

Per chiarire la differenza tra *directory* e motori veri e propri eseguiamo una ricerca utilizzando prima una *directory* e poi un motore di ricerca.

## Esempio di ricerca su directory

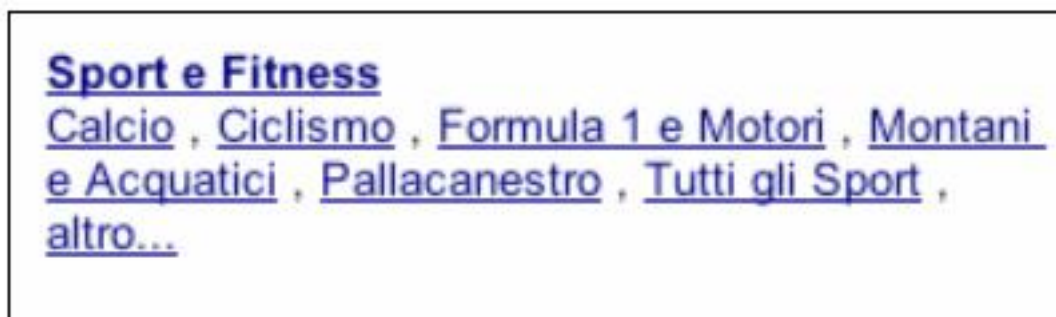
Supponiamo di voler ricercare informazioni sulla squadra di calcio dell'Inter. Utilizziamo la *directory* di Altavista (Figura 1) (<http://it.altavista.com/dir/>).



Pagina Iniziale della directory Altavista

Figura 1. Pagina Iniziale della *directory* Altavista.

Successivamente, tra le varie categorie elencate, clicchiamo su Sport e *Fitness* ottenendo una nuova pagina parte della quale è raffigurata in Figura 2.



Porzione della pagina Web di Altavista

Figura 2. Pagina iniziale della *directory* >Sport e *Fitness*.



Poi clicchiamo tra le varie sotto-categorie di Sport e *Fitness* su Calcio ottenendo una pagina parte della quale è raffigurata in Figura 3.



sottocategoria Calcio nella directory Altavista

Figura 3. Pagina iniziale della *directory* >Sport e *Fitness*>Calcio.

Clicchiamo adesso su Squadre....



sottocategoria Squadre nella directory Altavista

Figura 4. Pagina iniziale della *directory* >Sport e *Fitness*>Calcio>Squadre

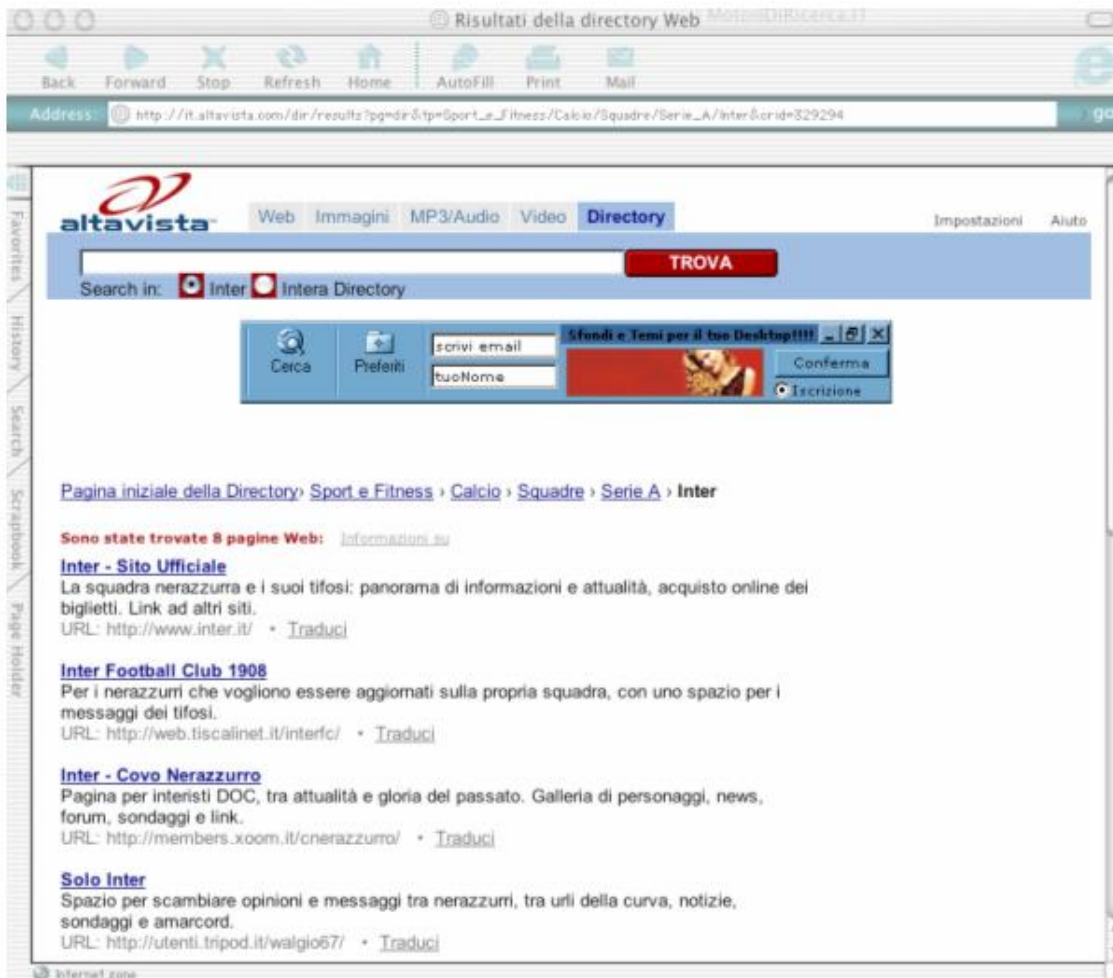
...e poi su Serie A...



sottocategoria Serie A nella directory Altavista

Figura 5. Pagina iniziale della *directory* >Sport e *Fitness*>Calcio>Squadre>Serie A

...e infine su Inter ottenendo una serie di *link* (risorse con relativo url) mostrati in Figura 6.



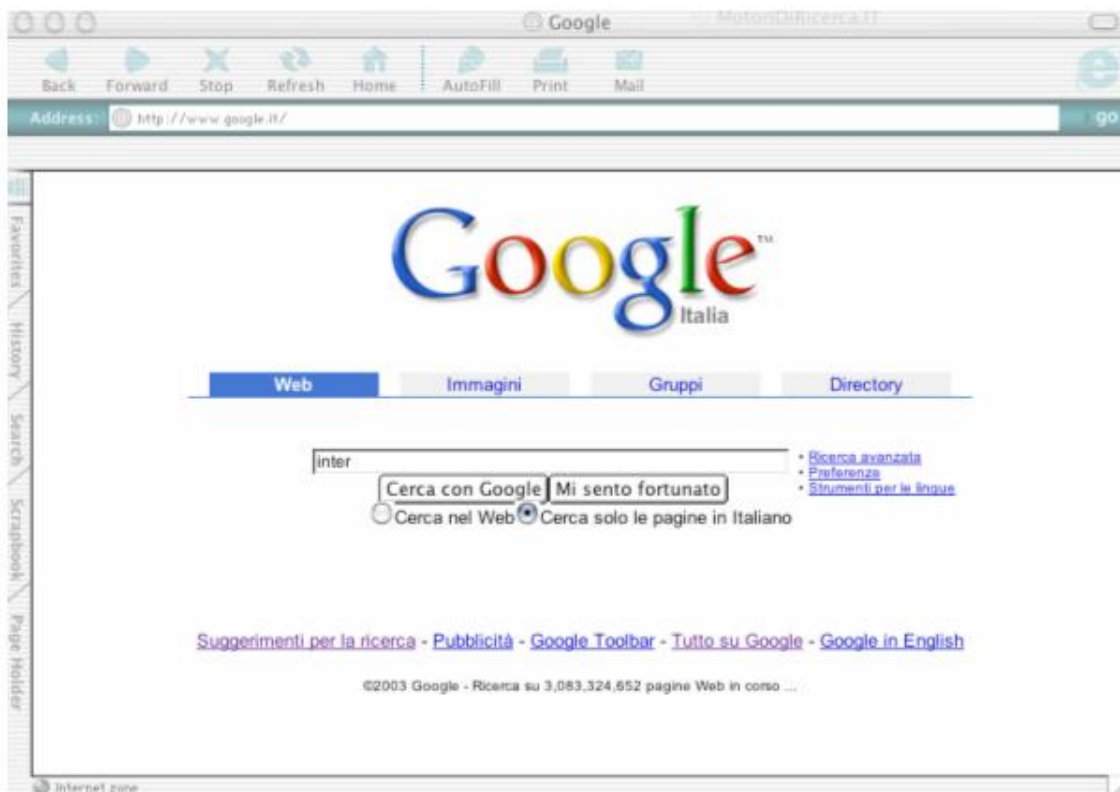
Pagina Web contenente il risultato della ricerca

Figura 6. Risultato della ricerca ...

## Esempio di ricerca su motore di ricerca

Cerchiamo ancora informazioni sull'Inter, però adesso utilizziamo un motore di ricerca vero e proprio. Per non far torto a nessuno questa volta usiamo il motore di ricerca *google* mostrato in Figura 7. (<http://www.google.it>).

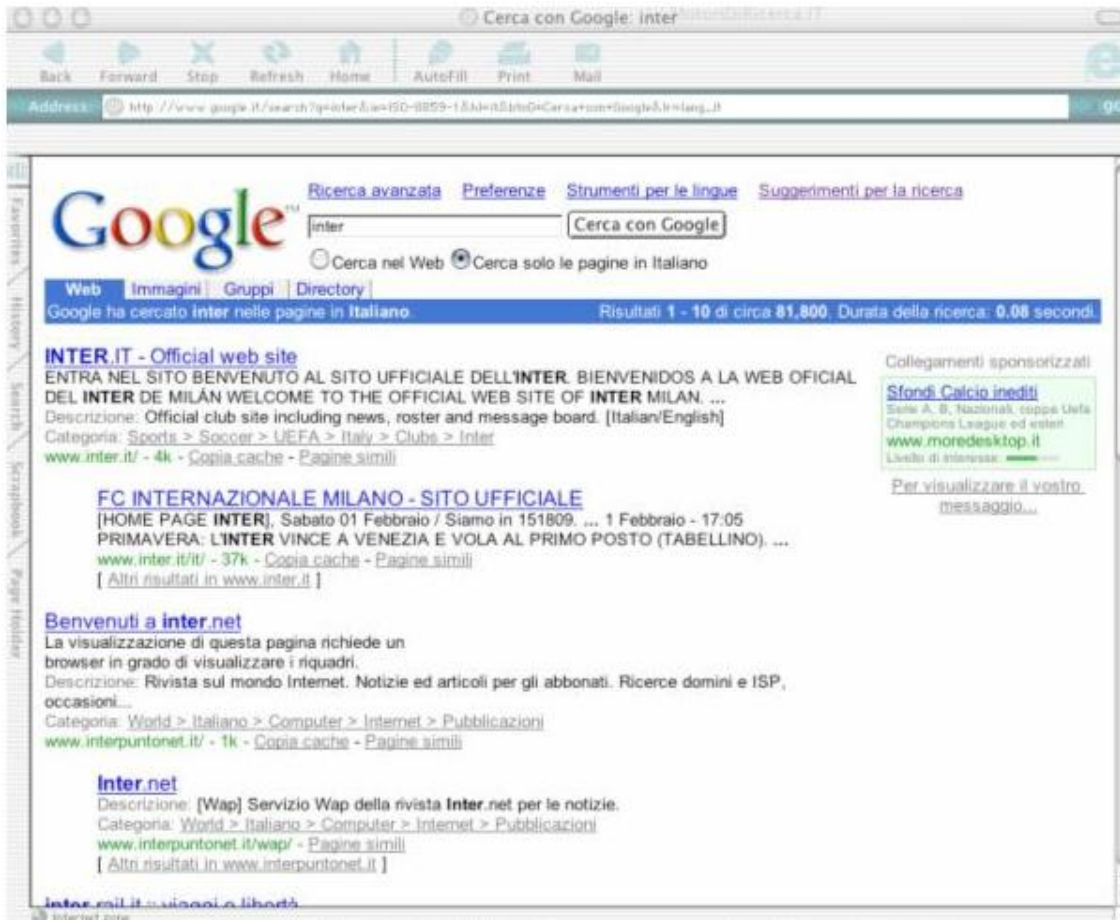
Per far ciò inseriamo la parola chiave *inter* e vediamo cosa ci restituisce *Google*...



Pagina Iniziale del motore di ricerca google

Figura 7. Pagina iniziale di *google*.

Otteniamo una lista (Figura 8) di 81800 *link* contenenti il termine *inter*. Siamo stati fortunati perché *Google* ci ha presentato i risultati in modo tale che tra i primi *link* proposti ci fossero proprio quelli riguardanti la squadra di calcio *Inter*. Non succede sempre così purtroppo...



Pagina Web contenente il risultato della ricerca effettuata con Google

Figura 8. Risultato della ricerca con parola chiave inter

Se scendiamo nella lista troviamo molti altri *link* contenenti la parola inter, ma non inerenti la squadra di calcio come ad esempio il *link* al Centro Inter-Bibliotecario dell'Università di Bologna (Figura 9).

[inter-rail.it :: viaggi e libertà](#)  
 ... Registrati gratuitamente Qui. Forum **Inter-Rail**, 6 recent Topics: · Primo reduce! ... Benvenuti su **Inter-Rail**, il sito dei liberi viaggiatori Ciao a tutti! ...  
[www.inter-rail.it/](#) - 70k - [Copia cache](#) - [Pagine simili](#)

[CIB Centro Inter-Bibliotecario - Universita' di Bologna : ...](#)  
 webmaster@cib.unibo.it, Alma Mater Studiorum Fondata nel 1088, Accesso ai Cataloghi.  
 Descrizione: Biblioteche, motori di ricerca su cataloghi di libri periodici riviste articoli di biblioteche dell'...  
 Categoria: [World](#) > [Italiano](#) > [Consultazione](#) > [Biblioteche](#)  
[www.cib.unibo.it/](#) - 7k - [Copia cache](#) - [Pagine simili](#)

[INTER EMOTIONAL SITE](#)  
 Foto **Inter**, **Inter**, **Inter** fc, foto 2002-03, fotointer, foto champions league, foto campionato, foto coppa italia, foto coppa uefa, foto mondiali, fotointer.com ...  
[digilander.libero.it/gointer/](#) - 42k - [Copia cache](#) - [Pagine simili](#)

[Forza Inter.it - Il portale dei tifosi interisti](#)  
**Inter** - Sito web non ufficiale dell'Internazionale Football Club Milano SpA. ... Ultima partita Lione-**INTER** 3-3 Prossima partita **INTER**-Rosenborg Tabellone. ...  
 Descrizione: Sito web non ufficiale dell'Internazionale Football Club Milano SpA. Passato, presente e futuro della...  
 Categoria: [World](#) > [Italiano](#) > [Sport](#) > [Calcio](#) > [Società](#) > [Inter](#)  
[www.forzainter.it/](#) - 23k - [Copia cache](#) - [Pagine simili](#)

Alcuni link non inerenti alla società di calcio Internazionale FC

Figura 9. Altri *link* non inerenti alla società di calcio Internazionale FC

Tipicamente, i siti di ricerca più importanti mettono a disposizione sia *directory* che motori di ricerca. La scelta tra i due non è sempre facile e dipende sostanzialmente dal tipo di risorsa che vogliamo ricercare e da quanto dettagliata è l'informazione che abbiamo a disposizione. In prima approssimazione, se l'informazione è molto dettagliata probabilmente converrà usare un motore di ricerca. Se l'informazione è generica, usando motori di ricerca rischiamo di trovare migliaia di *link* molti dei quali totalmente scorrelati da ciò che stiamo cercando.

Qui di seguito elenchiamo alcuni motori di ricerca (italiani e internazionali) e per ognuno di essi indichiamo quali motori di ricerca e quale sistema di *directory* vengono usati.

Nome	Motore	Directory
<i>Yahoo.com</i>	<i>Google</i>	proprietario
<i>Google</i>	proprietario	DMOZ
<i>Altavista.com</i>	proprietario	<i>Looksmart</i>
<i>Looksmart</i>	<i>Inktomi</i>	proprietario
<i>InWind</i>	Arianna	DMOZ
<i>Virgilio</i>	<i>Google</i>	proprietario
<i>KataWeb</i>	<i>Inktomi</i>	proprietario

Ad esempio *Google* utilizza un motore di ricerca proprietario (da lui stesso inventato) e una *directory* basata sul sistema DMOZ, mentre *Virgilio* utilizza il motore di ricerca fornito da *Google* e una *directory* proprietaria.

## Motori di ricerca

Un motore di ricerca si basa su diversi componenti:

- un programma che interroga periodicamente il *Web* per scaricare le pagine e catalogarle.
- Un *database* delle pagine catalogate.
- Una interfaccia di interrogazione.

### **Programma**

Il programma che cataloga le pagine *Web* è il cuore del motore di ricerca. È continuamente attivo e non fa altro che navigare su tutto il *Web*, scaricare pagine, estrarre le informazioni dalle pagine scaricate che ritiene utili per la loro classificazione e inserire queste informazioni nel *database*. Questo procedimento è molto complesso, richiede molto tempo ed è cruciale per tutte le altre fasi della ricerca.

### **Database**

Il *database* contiene l'informazione estratta dalle pagine *Web* esaminate dal programma. L'informazione contenuta nel *database* è organizzata in modo tale da permettere di ricercare in modo efficiente tutte le pagine che soddisfano i parametri di ricerca specificati dall'utente tramite l'interfaccia.

### **Interfaccia**

Quando un utente si collega a un motore di ricerca interagisce con la sua interfaccia per definire i parametri della ricerca. Solitamente la ricerca avviene sulla base di una combinazione di parole chiave inserite dall'utente.

### **Ricerche per parole chiave**

La ricerca più semplice è quella che consiste nel fornire al motore una sola parola chiave. Nell'esempio che abbiamo svolto all'inizio di questo approfondimento abbiamo fornito una sola parola chiave e la ricerca è andata a buon fine. Questo perché la parola chiave fornita specificava in modo preciso ciò che stavamo cercando. Non sempre è così, anzi. Nella maggioranza dei casi infatti una parola può essere utilizzata con diversi significati in diversi contesti.

Per questa ragione le interfacce di interrogazione dei motori di ricerca permettono di inserire più parole chiave legate tra di loro da un certo numero di operatori. Gli operatori più comunemente usati sono:

- *PrimaParola Or SecondaParola*: vengono ricercate le pagine che contengono almeno una tra *PrimaParola* e *SecondaParola*.
- *PrimaParola And SecondaParola*: vengono ricercate le pagine che contengono sia *PrimaParola* che *SecondaParola*.
- *PrimaParola But Not SecondaParola*: vengono ricercate le pagine che contengono *PrimaParola* e non contengono *SecondaParola*.
- *PrimaParola SecondaParola*: vengono ricercate le pagine che contengono la frase *PrimaParola SecondaParola*.
- *PrimaParola Near SecondaParola*: vengono ricercate le pagine che contengono *PrimaParola* e non troppo distante *SecondaParola*. È una restrizione dell'operatore *And*.
- *PrimaParola\**: vengono ricercate le pagine che contengono parole che iniziano con *PrimaParola*.



## Directory

Una *directory* è un sito *Internet* che vi permette di selezionare progressivamente attraverso il meccanismo gerarchico delle categorie e delle sotto-categorie i *link* alle risorse che vi interessano.

All'inizio della ricerca l'utente si trova di fronte ad un elenco di categorie molto generiche quali ad esempio politica, sport, tecnologia, eccetera. Selezionando una di queste categorie l'utente implicitamente elimina dalla ricerca tutti i siti che non appartengono alla categoria selezionata.

Procedendo con questo sistema di raffinamento progressivo l'utente arriva abbastanza velocemente a determinare una categoria specifica contenente *link* di suo interesse.

Questo tipo di ricerca è particolarmente utile quando l'utente non ha molto chiaro dall'inizio cosa sta cercando e ha bisogno di rendersi conto di che tipo di informazione può trovare sulla rete circa l'argomento di suo interesse. Spesso l'utente ha solo voglia di girellare su *Internet* senza un preciso percorso in mente. Anche in questo caso il meccanismo della ricerca per categorie può rivelarsi utile.

È cruciale in una *directory* l'organizzazione delle categorie e delle sotto categorie. Una buona *directory* deve mettere in grado l'utente di capire velocemente ad ogni passo quale categoria selezionare e al tempo stesso deve permettere di raggiungere la categoria finale con il minor numero possibile di scelte.

Solitamente le *directory* sono aggiornate a mano.

## Alcuni motori e directory

[www.yahoo.com](http://www.yahoo.com)

*Yahoo!* è la *directory* più grande del *Web*. Non è propriamente un motore di ricerca (anche se fornisce tale servizio appoggiandosi allo *spider* di *Google*), ma un catalogo di siti entro il quale si può navigare per categorie affinando progressivamente la scelta degli argomenti proposti. È stato fondato nel 1994 presso la *Stanford University* da *David Filo* e *Jerry Yang* quando erano ancora studenti di ingegneria elettronica.

Il loro scopo iniziale era quello di classificare siti a cui erano particolarmente interessati. Adesso *Yahoo!* è un motore popolarissimo in diretta competizione con *Altavista*. È stato il primo motore di ricerca ad essere pubblicizzato al di fuori di riviste specializzate di informatica al pari di qualsiasi altro prodotto commerciale.

*Yahoo!* ha fatto della ricerca per categorie il suo punto di forza. L'albero delle categorie proposto da *Yahoo!* è molto più esteso e meglio organizzato che altrove. *Yahoo!* è lo strumento di ricerca ideale per tutti coloro che iniziano una ricerca senza avere un'idea chiara di cosa stanno cercando e delle parole chiave che potrebbero descrivere ciò a cui sono interessati. È sicuramente lo strumento di ricerca più usato da utenti non esperti.

Il catalogo dei siti di *Yahoo!* viene continuamente aggiornato aggiungendo siti proposti dagli stessi autori. I siti proposti vengono analizzati, valutati ed eventualmente inseriti nel catalogo. Per proporre un sito all'attenzione dei *surfer* di *Yahoo!* (gli esperti che



valutano se inserire o meno il sito nel catalogo) basta andare sulla *home page* di *Yahoo!* e cliccare su *Proponi un sito con Yahoo! Express* (servizio a pagamento).

[www.google.com](http://www.google.com)

*Google* è una delle aziende con la tecnologia di ricerca più interessante. Sviluppato all'università di *Stanford* da *Sergey Brin* e *Larry Page*, due giovani di 26 anni laureati in scienze informatiche, ha debuttato *on-line* a fine settembre 1999 dopo oltre un anno di test.

*Google* al contrario di *Yahoo!* ha fatto della ricerca per parole chiave il suo punto di forza. La caratteristica di *Google*, è quella di selezionare i risultati di ricerca valutando l'importanza di ogni pagina *Web* con metodi matematici, in base ad un controllo di oltre 500 milioni di variabili e di 2 miliardi di termini. Questa tecnologia, chiamata *PageRank* (brevettata da *Google*), controlla non solo il contenuto della pagina *Web*, ma verifica anche altri eventuali siti che hanno un *link* verso la pagina: in base alla quantità ed al tipo di *link*, la pagina riceve una valutazione più o meno alta.

Altre caratteristiche di *Google* sono la sua velocità di esecuzione, con un tempo medio dichiarato per ogni ricerca di 0,29 secondi, e la grandezza del suo archivio che ha oltre 3 miliardi di pagine.

La convinzione di fornire risultati rilevanti, è confermata dal bottone *I'm feeling lucky* (mi sento fortunato), attraverso il quale l'utente può essere direttamente portato sul primo sito che *Google* avrebbe indicato a fronte dei termini indicati.

L'interfaccia di *Google* è volutamente essenziale e priva di pubblicità.

[www.virgilio.it](http://www.virgilio.it)

*Virgilio* ha tutti i numeri per risultare una delle più importanti *Guide Internet*. Censisce circa 100.000 siti italiani con l'aggiunta di 150 nuovi siti ogni giorno, ed offre una vastissima gamma di servizi funzionali e di orientamento, tali da considerarlo come un vero punto di partenza per ogni ricerca sulla rete, e non solo in Italia.

Il cuore di *Virgilio* è la sua *directory*, ossia un elenco di siti suddiviso in base alla loro tipologia, e strutturato in macro categorie a loro volta distinte in gruppi più analitici.

La registrazione dei siti è effettuata manualmente da uno staff interno a *Virgilio* (servizio a pagamento) e ciò permette una classificazione efficiente e coerente con la suddivisione in categorie. Sono censiti solo i siti in lingua italiana, con l'unica esclusione di quelli i cui contenuti sono ritenuti non idonei, ad esempio quelli pornografici, violenti, eccetera.

[www.overture.com](http://www.overture.com)

*Overture* è nata a fine 1997 a *Pasadena* in California ed è arrivata *on-line* il 21 febbraio 1998. Il suo fondatore, *Bill Gross*, già creatore nel 1988 del programma *Lotus Magellan*. Il CEO è *Jeffrey Brown*, in passato cofondatore di *City Search*.

*Overture* è sempre stato un motore di ricerca piuttosto originale rispetto ai più noti siti del settore, sia per caratteristiche che per strategia.

Il modello di *business* di *Overture* è quello di far pagare il posizionamento in testa ai

risultati delle ricerche in base alle visite effettivamente generate. Ad esempio, un'azienda che vende vino *on-line* può fare in modo che il suo sito figuri in testa alle ricerche effettuate dagli utenti con i termini vino, *wine*, vini italiani, eccetera. Per ottenere le posizioni di testa, l'azienda si impegnerà a pagare a *Overture* una cifra per ogni click effettuato, in base ad una specie di asta. Se infatti un produttore di vini concorrente volesse figurare in posizione predominante, dovrebbe impegnarsi a pagare una cifra superiore.

I risultati delle ricerche forniti da *Overture* sono quindi piuttosto rilevanti, evidenziando siti di aziende certamente affini con le richieste fatte dall'utente, il quale peraltro ha ben evidenziate le posizioni a pagamento e l'ammontare che il destinatario è disposto a pagare a *Overture* per la sua visita.

La qualità delle risposte, è anche assicurata dall'utilizzo della tecnologia *Inktomi* per la gestione dei risultati di ricerca successivi a quelli a pagamento. Inoltre *Overture* non sembra voler mettere i vestiti del portal come hanno fatto tutti i principali motori di ricerca, per cui non vengono presentati altri contenuti sul sito se non quelli riguardanti le ricerche effettuate dall'utente.

Oltre ad implementare per primo il sistema di posizionamento a pagamento, *Overture* è stato anche il primo motore di ricerca ad introdurre un *Affiliate Program*, ossia un accordo con altri siti, i quali possono esporre un box di ricerca di *Overture* ricavandone una commissione per ogni ricerca effettuata dai loro utenti.

## Motori tematici

I motori di ricerca di cui abbiamo parlato finora sono detti motori generalisti. Ciò significa che forniscono informazioni su risorse *Internet* di qualsiasi tipo e su qualsiasi argomento.

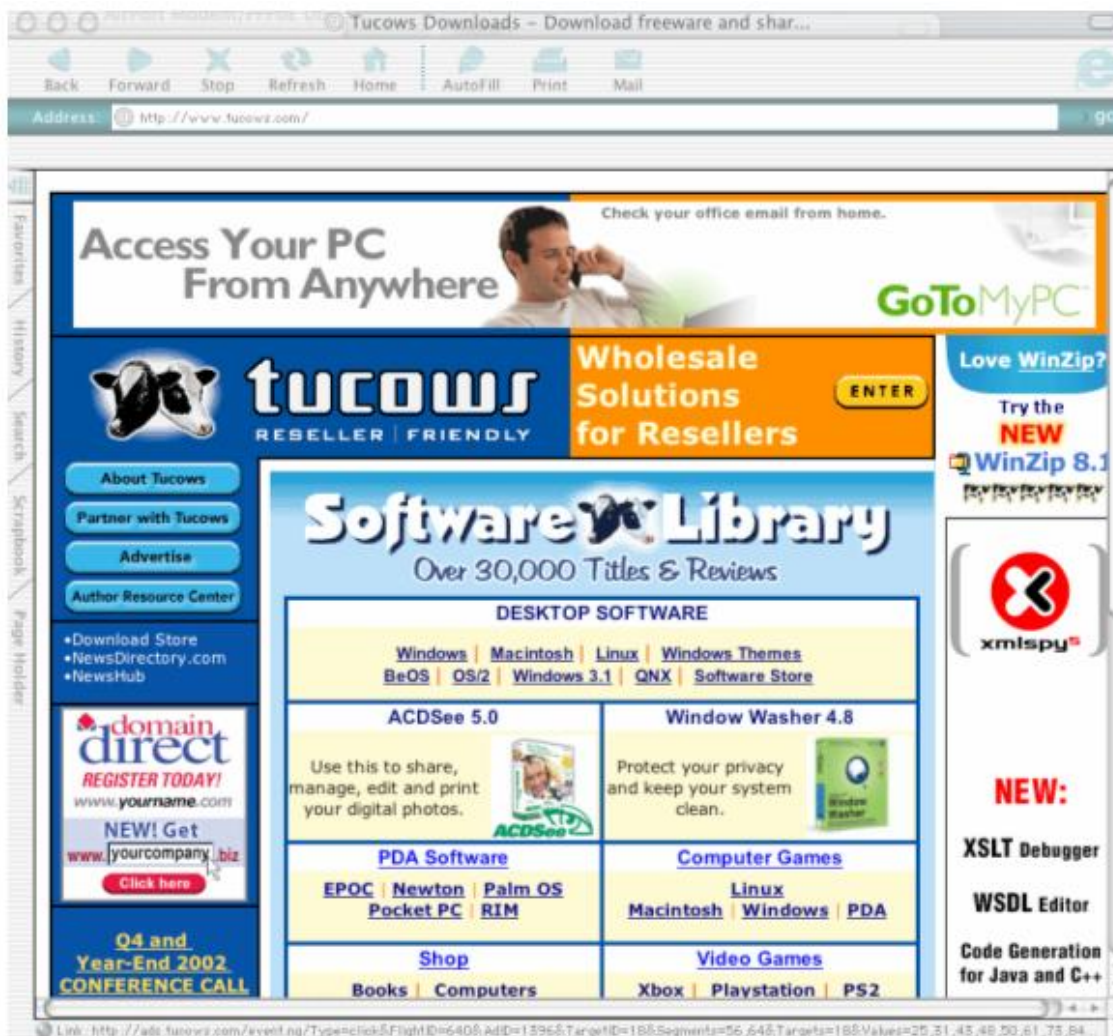
Esistono però anche strumenti di ricerca tematici o finalizzati che sono specializzati nelle ricerche di risorse *Internet* su determinati argomenti.

Ci sono motori ad esempio specializzati nella ricerca di *software*, di *file* musicali (ad esempio *file* in formato mp3), di annunci immobiliari, di offerte di lavoro e di molto altro ancora.

Questi motori solitamente riescono a fornire risultati più accurati sullo specifico tema trattato rispetto ai motori generalisti.

### Un Esempio: *Tucows*

*Tucows* può essere considerato a tutti gli effetti un motore di ricerca per scovare nella rete il *software* che vogliamo. *Tucows* è l'acronimo di *The Ultimate Collection Of Windows Software*, ma è anche un gioco di parole, ovvero è l'equivalente della pronuncia inglese di *two cows* che significa due mucche; ecco il motivo dell'icona raffigurante per l'appunto due mucche simbolo ufficiale del sito [www.tucows.com](http://www.tucows.com), vedi Figura 10.



Pagina Iniziale del sito Web twocows

Figura 10. Pagina iniziale del sito *Tucows*.

In questo sito è possibile ricercare *software* sia *freeware* che a pagamento per ogni piattaforma (tra le principali citiamo *unix*, *linux*, *windows* e *macintosh*).

## Meta-Motori

Infine, senza voler scendere in dettaglio vorremmo menzionare l'esistenza dei meta-motori. Un meta-motore non è altro che un sito che si appoggia sui servizi forniti da un certo numero di motori di ricerca. L'utente usa l'interfaccia del meta-motore per impostare i suoi parametri di ricerca e riceve in *output* dal meta-motore il risultato ottenuto mediante la consultazione di più motori di ricerca.

Questo modo di ricercare informazioni può risultare più efficace rispetto alla consultazione di un singolo motore di ricerca nei casi in cui la nostra ricerca sia particolarmente difficile da portare a termine. I singoli motori di ricerca infatti coprono ciascuno non più del 35-40% di *Internet*. Usare più motori di ricerca

contemporaneamente può quindi risultare utile.

Url di alcuni meta-motori:

- [www.thebighub.com](http://www.thebighub.com)
- [www.cerca.com](http://www.cerca.com)
- [www.dogpile.com](http://www.dogpile.com)
- [www.highway61.com](http://www.highway61.com)
- [www.metacrawler.com](http://www.metacrawler.com)

**Kartoo** ([www.kartoo.com](http://www.kartoo.com))

*Kartoo* è un meta-motore. La caratteristica principale di *Kartoo* è quella di rappresentare graficamente i risultati della ricerca sotto forma di una o più mappe (Figura 11). Inserite le parole chiave, *Kartoo* effettua la ricerca sui principali motori di ricerca in circolazione, mette insieme i risultati, e li riveste graficamente affinché possano essere utilizzati in modo semplice e intuitivo dall'utente.



Risultato della ricerca con parola chiave 'batistuta' sul motore di ricerca Kartoo

Figura 11. Risultato della ricerca con parola chiave batistuta sul motore di ricerca *Kartoo*.

Il risultato della ricerca è visualizzato sotto forma di mappa di collegamenti.

La mappa restituita da *Kartoo* è composta da diverse sezioni.

Nella sezione circolare di Figura 11 abbiamo il risultato vero e proprio della ricerca composto da un insieme di cerchi (nodi) di diverse dimensioni e da un insieme di collegamenti (*topics*) che collegano un cerchio ad un altro.

Ogni cerchio rappresenta un sito *Internet* e la sua dimensione è proporzionale alla suo grado di importanza all'interno della ricerca. Cerchi più grandi rappresentano pagine molto significative all'interno delle quali è più probabile trovare ciò che stiamo cercando. I collegamenti tra un nodo e un altro sono etichettati con una o più parole chiave che possono essere aggiunte alla ricerca (cliccandoci sopra) in fasi successive. La ricerca avviene quindi per raffinamenti successivi ottenuti inserendo o escludendo una sequenza di parole chiave (suggerite dalla mappa o inserite dall'utente).

Inoltre i siti più significativi vengono visualizzati al centro della mappa;

In ogni momento della ricerca è possibile **inserire il proprio sito nei motori di ricerca**.

## Inserire il proprio sito nei motori di ricerca

Finora abbiamo trattato il problema di ricercare l'url di risorse in *Internet* fornendone una descrizione più o meno dettagliata attraverso vari meccanismi (parole chiave, categorie di appartenenza). Supponiamo adesso invece di voler affrontare il problema inverso e cioè come far sì che il maggior numero di utenti possibile raggiunga il nostro sito *Internet* per effetto di una ricerca su motori o *directory*.

Questa problematica diventa di fondamentale importanza nel momento in cui, ad esempio, si intraprenda una attività di commercio elettronico e si voglia quindi pubblicizzare il nostro prodotto:

- Come è possibile inserire il nostro sito in una *directory*?
- Come è possibile far sì che il nostro sito appaia nelle prime posizioni nella lista di url ottenuti da un determinato motore di ricerca?

Immaginatevi di voler ad esempio vendere via *Internet* un prodotto tipico della vostra zona geografica di provenienza. Come è possibile pubblicizzare la vostra iniziativa? Ovvero far conoscere a più persone possibili la vostra offerta commerciale?

Il problema non è nuovo. Chiunque voglia entrare nel mondo del commercio ha bisogno di pubblicità. Su *Internet* sono disponibili però strumenti specifici per acquisire notorietà. Rimangono validi gli strumenti pubblicitari classici: televisione, radio, giornali, manifesti, ma a questi se ne affianca uno nuovo come già avrete immaginato, ovvero *Internet* stessa tramite l'utilizzo dei motori e delle *directory*.

Non vogliamo entrare nei dettagli di questa problematica ampia e complessa, ma desideriamo comunque chiarire alcuni concetti di base:

- La prima cosa da dire è che un sito dotato di grafica eccezionale e di una struttura ben organizzata e completa, ma non inserito in nessun motore di ricerca rischia di non riscuotere il successo che merita!
- I motori di ricerca veri e propri navigano in *Internet* autonomamente cercando nuovi siti da catalogare senza bisogno che nessuno suggerisca loro di inserire un certo sito nel *database*. Le *directory* invece, al contrario dei motori di ricerca, svolgono una esplorazione del *Web* limitata in quanto tale attività viene svolta tipicamente da esseri umani (*surfer*) senza l'impiego di strumenti informatici automatizzati. Il metodo più efficace quindi per inserire un sito in una *directory* è senza dubbio quello di proporre il

proprio sito per la catalogazione invece di sperare che esso venga notato e catalogato dalla *directory*.

- Spesso registrare un sito in una *directory* costa denaro, soprattutto se la *directory* in questione è una delle più importanti e diffuse.
- Quando si effettua una registrazione presso un motore o una *directory* occorre fornire informazioni sul sito che vogliamo registrare. Solitamente viene richiesto all'autore del sito di fornire parole chiave oppure di scegliere una o più categorie nelle quali inserire il sito. Lo stesso linguaggio HTML (linguaggio per scrivere pagine *Web*) mette a disposizione costrutti sintattici specifici per poter inserire le parole chiave in base alle quali poi la ricerca verrà effettuata.
- Alcuni motori (esempio *Google*) oltre ad associare ai siti le relative parole chiave operano anche utilizzando il cosiddetto *ranking*. Il *ranking*, come abbiamo già accennato, è un meccanismo di valutazione dei siti in base alla loro importanza alla loro visibilità e al numero di visite ricevute. I siti con un *ranking* più alto saranno visualizzati nelle ricerche in posizioni migliori.

## Web nascosto o dinamico

Non tutte le pagine che visualizziamo sul nostro *browser* sono catalogate nei *database* dei motori di ricerche. Alcune di esse infatti non esistono! Per essere più precisi esistono solo per un certo lasso di tempo (il tempo durante il quale sono visualizzate sul vostro *browser*) solo perché la vostra interazione con un sito *Web* le ha generate.

Facciamo un esempio per chiarire ciò che abbiamo appena detto. Supponiamo di collegarci ad un sito che gestisce la compravendita di autovetture nuove ed usate. Il sito vi permetterà ad esempio di specificare marca e modello dell'auto che state cercando. Fatto questo il sito genererà dinamicamente una pagina (il suo codice html) contenente le offerte che fanno al caso vostro. Quella pagina è frutto di una interazione tra noi e il sito su cui stiamo navigando e viene generata ad-hoc da un programma che gira sul *server* dove il sito è pubblicato. Quella pagina ovviamente non è catalogata in nessun motore di ricerca.



## Cliente di posta elettronica

Luciano Margara

12.2.2 (Usare tool e utilità Internet quali: e-mail, browser, motori di ricerca, newsgroup, list server, chat, trasferimento di file)

### Di cosa abbiamo bisogno per usare la posta elettronica?

#### Di cosa abbiamo bisogno per usare la posta elettronica ?

- Accesso a Internet (ISP).
- Account di posta elettronica
- Cliente di posta (programma per editare messaggi)

Di cosa abbiamo bisogno per usare la posta elettronica?

L'obiettivo di questo approfondimento è imparare a configurare un cliente di posta elettronica. Come abbiamo già detto nella parte introduttiva di questo modulo, per utilizzare il servizio di posta elettronica abbiamo sostanzialmente bisogno di tre cose: un accesso ad *Internet*, di solito fornito da un *Internet Service Provider*, un *account* di posta elettronica e un cliente di posta, che è un programma che serve per editare, cioè comporre, spedire e ricevere messaggi di posta elettronica.

### Accesso a Internet



## Accesso a Internet

- Per spedire e ricevere messaggi di posta elettronica occorre essere connessi a Internet.
- La connessione a Internet viene fornita da un Internet Service Provider
- Esempio: Tiscali, Wind, Infostrada, ...

Accesso a Internet

Per utilizzare il servizio di posta elettronica abbiamo bisogno di avere un accesso ad *Internet*. L'accesso ad *Internet* solitamente ci viene fornito da un *Internet Service Provider*, al quale ci connettiamo col nostro modem. Esempi di *Internet Service Provider* sono: Tiscali, *Wind*, Infostrada, ma ce ne sono molti altri.

### **Account di posta elettronica**

# Account di Posta Elettronica

- Casella di posta elettronica
- Indirizzo di posta elettronica

Account di posta elettronica

In sostanza un *account* di posta elettronica consiste di una casella di posta elettronica, che è uno spazio disco ospitato da un *host* di rete. In questo spazio disco, che è appunto la nostra casella di posta elettronica, noi memorizziamo i messaggi che ci vengono inviati e quando abbiamo bisogno di leggere i nostri messaggi ci connettiamo ad *Internet* e li scarichiamo dalla nostra casella di posta elettronica. Inoltre abbiamo bisogno di un indirizzo, che è parte integrante del nostro *account* di posta elettronica. L'indirizzo di posta elettronica viene utilizzato da tutti quelli che vogliono spedirci un messaggio.

## Configurare un cliente di posta 1

## Configurare un cliente di posta

- Il cliente di posta ha bisogno di essere configurato per poter funzionare.
- Bisogna (tra le altre cose) inserire i dati del nostro account di posta nel programma.

Configurare un cliente di posta 1

Quando utilizziamo il nostro cliente di posta elettronica per la prima volta abbiamo bisogno di configurarlo. La procedura di configurazione di un cliente di posta è sostanzialmente una procedura che informa il cliente di posta dei dati relativi al nostro *account*. Quindi abbiamo bisogno di entrare nelle preferenze del programma e di inserire tutti i dati relativi al nostro *account*; non solo quelli, ma principalmente quelli perché sono i più importanti.

## Configurare un cliente di posta 2

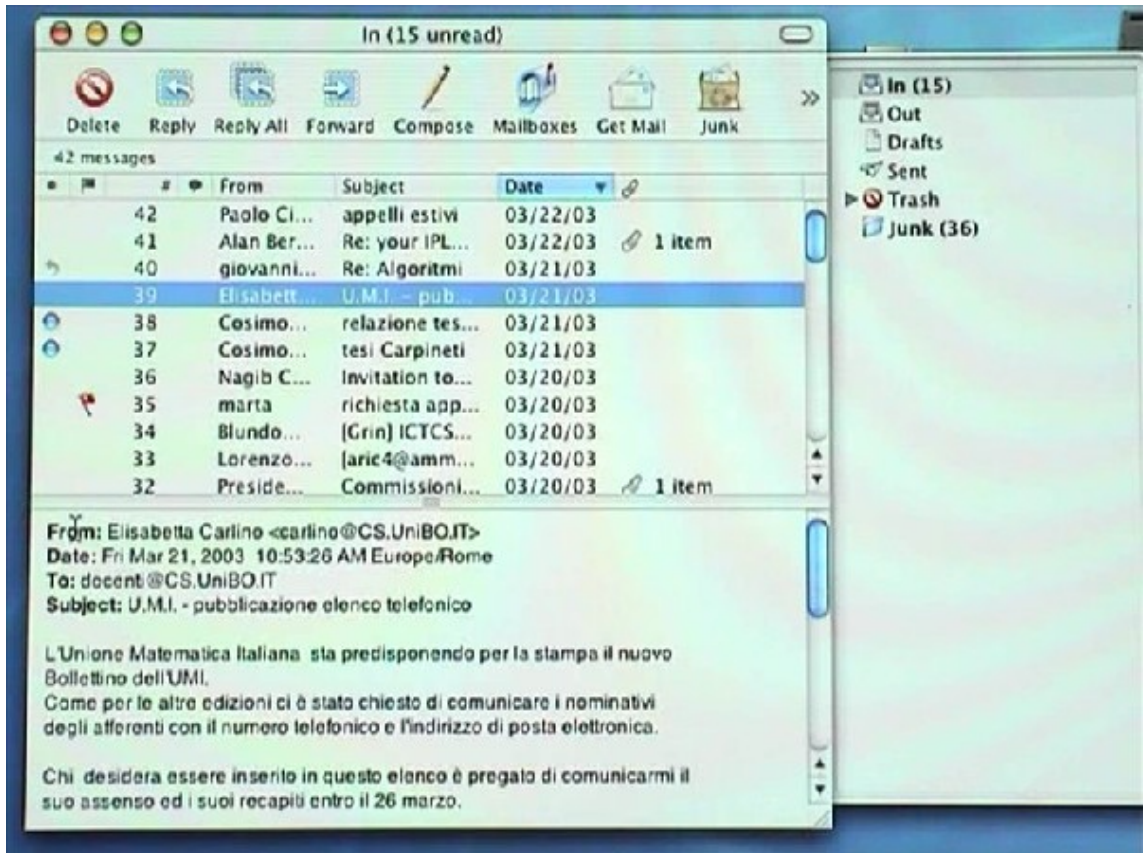
- In questo approfondimento impareremo a configurare i parametri principali di un client di posta elettronica.
- Utilizzeremo il client di posta *MAIL* fornito con *Mac OS X*.
- Non perdiamo di generalità.

Configurare un cliente di posta 2

In questo approfondimento impareremo per l'appunto a configurare i parametri principali di un cliente di posta elettronica. Per far questo utilizzeremo un cliente che si chiama *Mail* che è fornito col sistema operativo *Mac OS X*, che è la versione più recente del sistema operativo montato su *computer Apple*. Facendo questo non perderemo di generalità perché ciò che impareremo a fare sul cliente di posta *Mail* in realtà dovremo farlo comunque anche se utilizzassimo un qualsiasi altro cliente di posta elettronica, per esempio uno che gira sotto il sistema operativo *Windows*.

## Esempio di cliente di posta 1

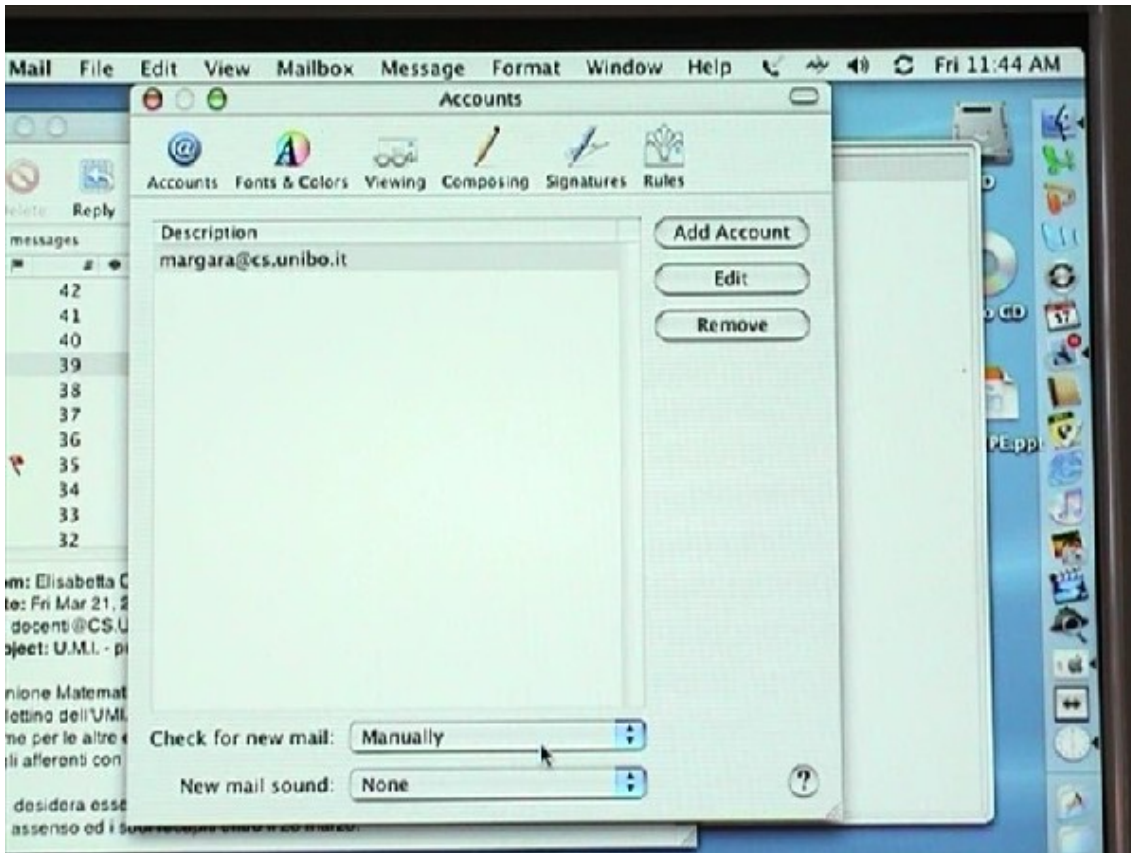




Esempio di cliente di posta 1

Una volta avviato il nostro cliente di posta elettronica *Mail* ci compare una schermata di questo tipo: in questa schermata abbiamo tre zone distinte. La prima, che è questa sulla destra, non è altro che un elenco delle *mailbox* che abbiamo a disposizione, relative al nostro *account* di posta elettronica: abbiamo una *mailbox IN* e una *mailbox OUT* che sono le due *mailbox* principali. Nella prima abbiamo memorizzati i messaggi che ci sono arrivati e che abbiamo scaricato dal nostro *Internet Service Provider*, questi messaggi sono elencati in questa zona. Notiamo che in questa colonna è indicato il mittente, il soggetto e la data di spedizione di ogni messaggio. Qui abbiamo selezionato il messaggio numero 39. Nella finestra indicata dal puntatore del mouse abbiamo i dettagli del messaggio: il mittente, la data, il destinatario, il soggetto e il messaggio.

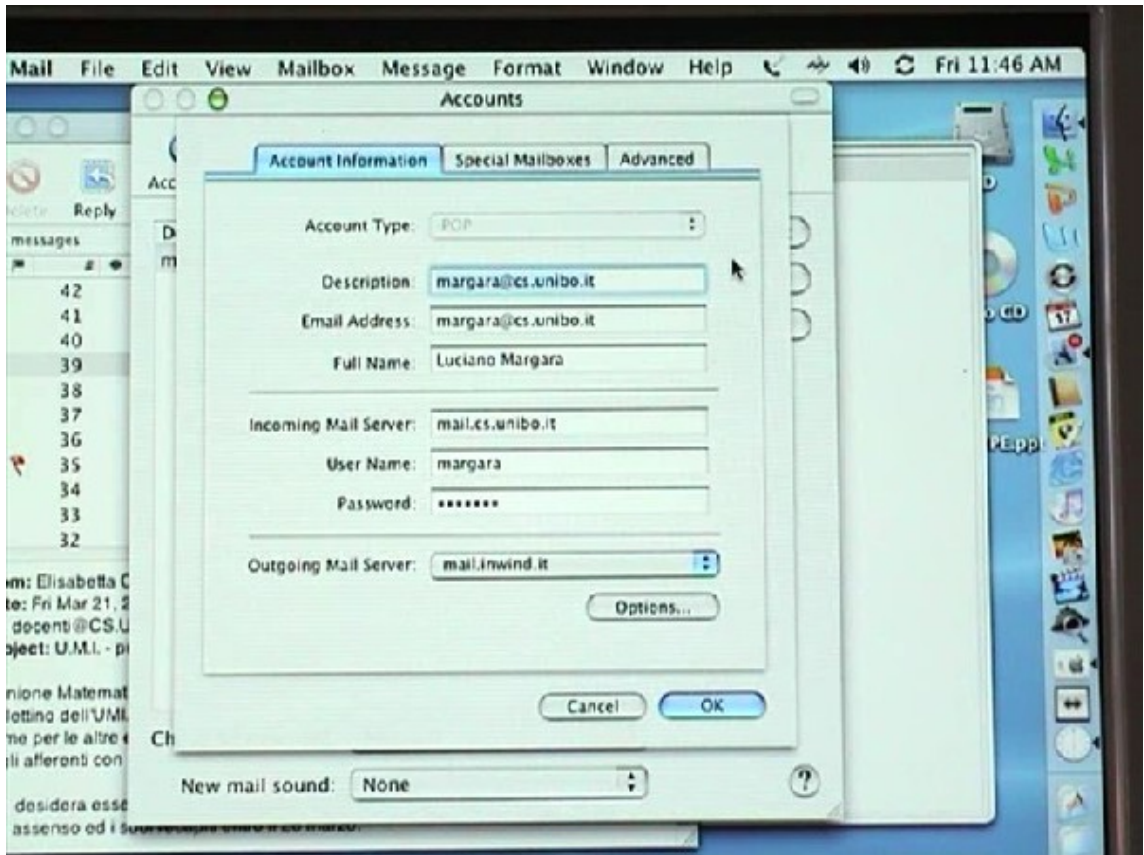
## Add account



Add account

Attiviamo adesso il pannello preferenze del nostro cliente di posta (in inglese *preferences*). In questo pannello abbiamo l'elenco degli *account* di posta già configurati: ce n'è uno solo, che si chiama *margara@cs.unibo.it*, che è stato configurato in precedenza, nel momento in cui per la prima volta è stato attivato il nostro *mailer*. Agendo sui pulsanti posizionati sulla parte superiore della finestra possiamo procedere alla configurazione del nostro cliente di posta relativamente agli *account*, ai *font*, ai colori, a come vedere i messaggi, a quali opzioni nel comporre i messaggi, a eventuali firme che vogliamo apporre ai messaggi. Possiamo anche configurare parametri relativi alle regole che vogliamo che siano attivate nel momento in cui riceviamo un messaggio, saremo più precisi in seguito. I primi due parametri che il pannello di preferenze ci chiede di configurare sono: *check for new mail* e *new mail sound*. Il parametro *check for new mail*, sostanzialmente, ci permette di scegliere se scaricare i messaggi di posta elettronica dal *Mail Server* del nostro *Internet Service Provider* manualmente, cioè nel momento in cui decidiamo noi, oppure se dire al nostro programma di posta elettronica di controllare se ci sono nuove *e-mail* ad intervalli di tempo regolari. Il parametro *new mail sound*, semplicemente, ci permette di specificare quale suono il nostro programma di posta elettronica deve emettere nel momento in cui riceviamo un nuovo messaggio di posta elettronica. Questo è il modo con cui il nostro programma di posta elettronica ci avverte che abbiamo un nuovo messaggio che ci attende.

## Account information 1

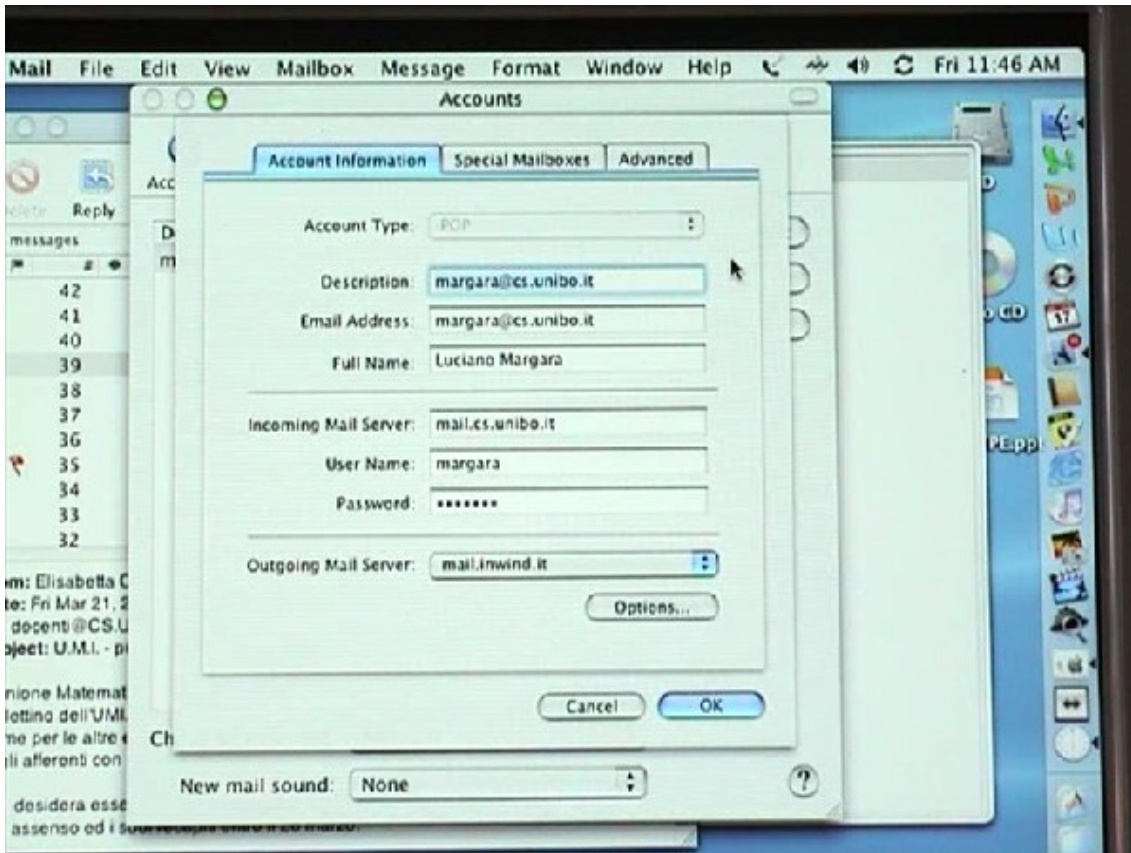


Account information 1

Prendiamo, adesso, in esame in dettaglio i parametri che dobbiamo configurare relativamente al nostro *account*. Qui abbiamo in elenco un unico *account* il cui nome è: `margara@cs.unibo.it`. Ho scelto questo nome simbolico per il mio *account* in quanto questo è l'indirizzo di posta elettronica relativo a questo *account*, ma in realtà avrei potuto scegliere un qualsiasi altro nome simbolico che identificasse univocamente il mio *account*. Se clicchiamo su *edit*, a questo punto, ci compare una finestra nella quale noi possiamo specificare diverse informazioni. Il primo parametro da specificare quando si configura un *account* è il suo nome, un nome simbolico qualunque; in questo caso io ho utilizzato l'indirizzo di posta elettronica associato al mio *account*. In questo campo dobbiamo inserire l'indirizzo di posta elettronica. Quale indirizzo di posta elettronica che ci ha fornito il nostro *Internet Service Provider* nel momento in cui gli abbiamo richiesto un *account* di posta elettronica. Se vogliamo utilizzare il servizio di posta elettronica, la cosa più facile da fare è collegarsi ad *Internet* al sito di uno degli *Internet Service Provider* che forniscono questo servizio e registrarsi. Nel momento in cui ci registriamo ci viene attribuito un indirizzo di posta elettronica e ci vengono comunicati una serie di altri dati importanti per configurare il nostro *account*. Qui, in questa finestra, dobbiamo inserire l'indirizzo di posta elettronica che ci è stato attribuito dall'*Internet Service Provider*. Nella finestra sottostante, questa evidenziata, dobbiamo inserire il nostro nome completo: questo sarà il nome che comparirà nel campo mittente nei messaggi che noi spediamo.

## Account information 2

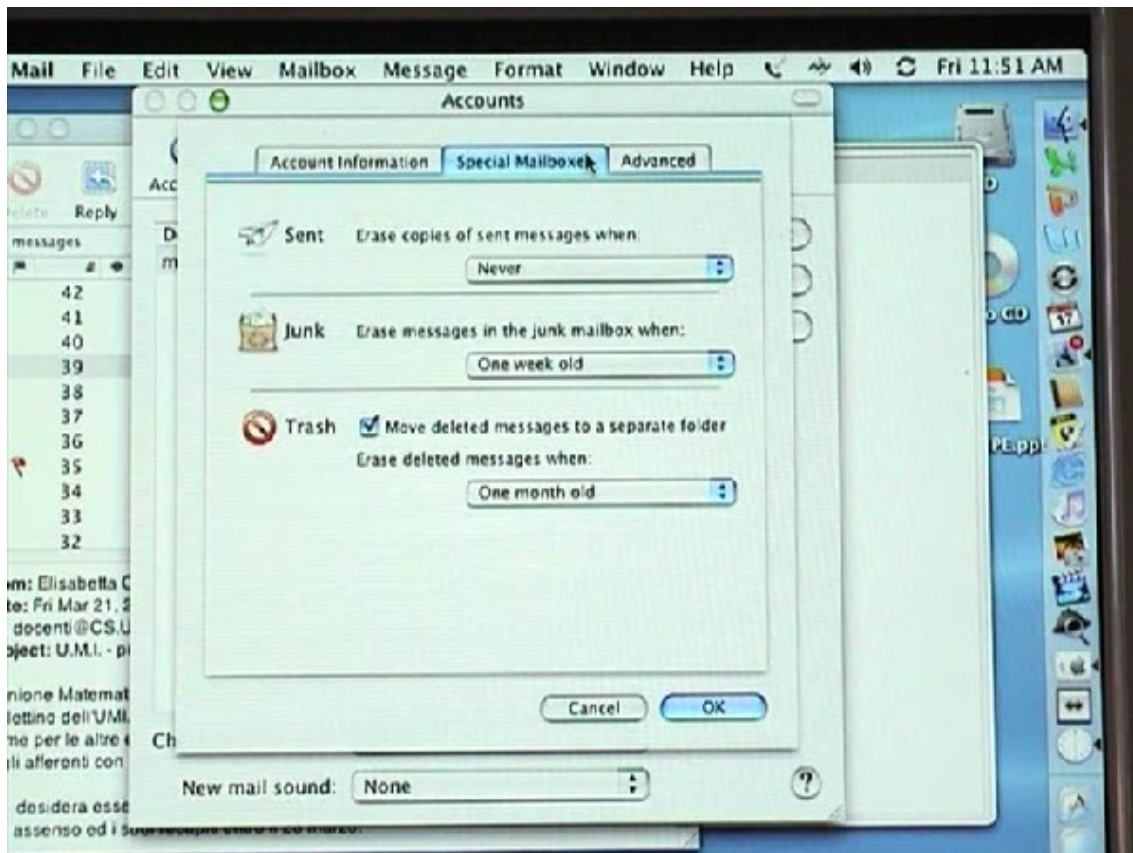




Account information 2

L'*Internet Service Provider* che ci ha fornito l'indirizzo di posta elettronica possiede, per gestire la posta elettronica, sostanzialmente due macchine, chiamate *Mail Server*. Il primo si chiama *Incoming Mail Server*, è un *host* di rete destinato a memorizzare i messaggi che arrivano alla nostra *mailbox*. Quindi `mail.cs.unibo.it` non è altro che l'indirizzo di un *Mail Server*, indirizzo simbolico, che ci è stato fornito al momento della registrazione. Questa è una informazione che ci deve dare il nostro *Internet Service Provider* nel momento in cui noi gli chiediamo di utilizzare il servizio di posta elettronica presso di lui. Quando ci registriamo, dobbiamo anche inserire uno *user name*, che poi verrà riportato in questo campo, per esempio il cognome, e la *password* che abbiamo fornito quando ci siamo registrati. *User name* e *password* servono, nel momento in cui vogliamo scaricare la nostra posta elettronica, al *Mail Server*, il quale dallo *user name* e dalla *password* capisce che siamo veramente noi le persone autorizzate a leggere quella determinata posta elettronica. Quindi ci serve in fase di autenticazione, quando ci colleghiamo per scaricare la nostra posta elettronica. C'è un altro *Mail Server* che entra in gioco nella gestione della posta elettronica, che si chiama *Outgoing Mail Server*. L'*Outgoing Mail Server* è una macchina del nostro *Internet Service Provider* che si occupa di prendere i messaggi che noi vogliamo spedire e instradarli sulla rete, verso il destinatario. Anche il nome di questo *Mail Server* ci viene fornito dall'*Internet Service Provider* che ci fornisce il servizio di posta elettronica; altra informazione che dobbiamo ricevere nel momento in cui ci registriamo.

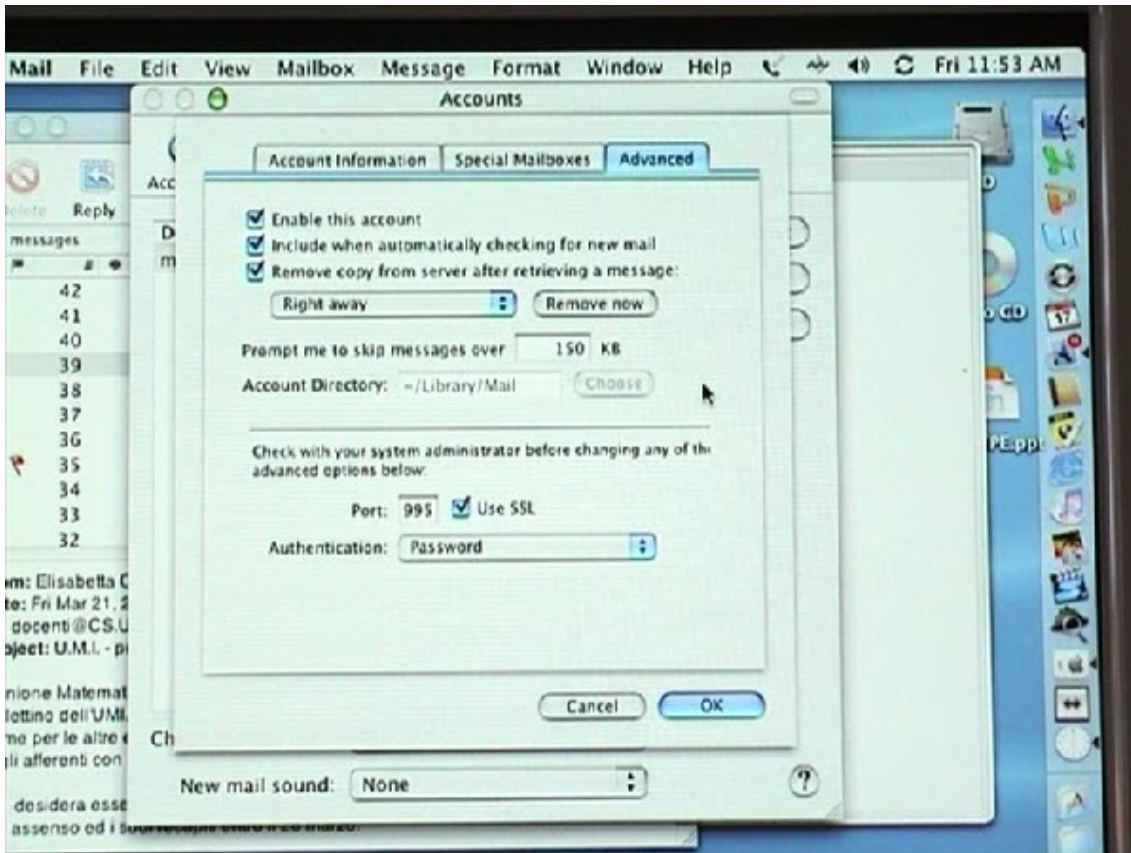
## Special Mailboxes



Special Mailboxes

Il programma *Mail* che stiamo descrivendo memorizza in 3 *mailbox* separate: i messaggi che noi abbiamo spedito (*sent mail*), i messaggi che noi riteniamo *junk mail*, ovvero *mail* indesiderate ricevute da mittenti che noi non conosciamo (per esempio messaggi pubblicitari) e *trash mail*, ovvero le *mail* che noi abbiamo deciso di cancellare. In questo pannello noi possiamo decidere diverse cose relativamente ai messaggi spediti, alla *junk mail* e alla *trash mail*. Sostanzialmente possiamo stabilire quando cancellarli; in questo menù a tendina, per esempio, possiamo dire al nostro programma di posta elettronica: cancella i messaggi spediti dopo un giorno, dopo una settimana, dopo un mese. In questo menù a tendina, invece, possiamo dire: cancella la *junk mail* dopo un giorno, dopo una settimana, dopo un mese. Mentre in quest'ultimo menù a tendina, possiamo dire: cancella i messaggi fisicamente dal nostro computer dopo un giorno, o una settimana, o dopo un mese. Il procedimento di cancellazione di un messaggio avviene in due fasi. Nella prima fase trasferiamo il messaggio nel *trash* (nel cestino), ma ancora il messaggio si può leggere e si può di nuovo trasferire in una delle altre *mailbox*. Quando invece noi lo cancelliamo, usando un apposito pulsante, questo *mail* viene eliminato dal nostro *hard disk* e non può essere più letto.

## Advanced 1

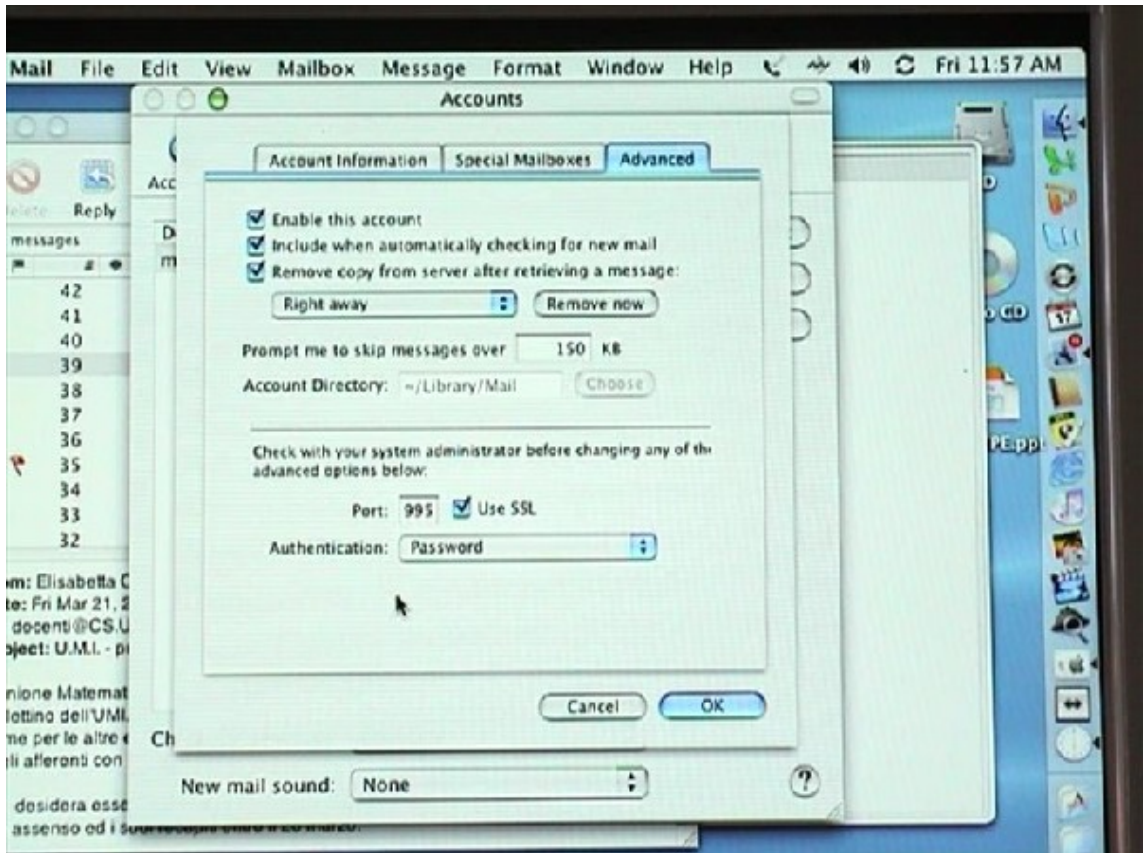


Advanced 1

Questo pannello di configurazione, chiamato avanzato, è diviso in due parti. Nella prima parte possiamo agire su queste caselle per configurare il nostro cliente di posta elettronica, in particolare possiamo decidere se vogliamo attivare l'*account* che stiamo configurando. Questa è una scelta molto utile perché una persona può avere più *account*. Alcuni magari sono vecchi e vogliamo disattivarli, alcuni non vogliamo utilizzarli e li disattiviamo, alcuni vogliamo utilizzarli e allora li attiviamo ponendo un *flag*, quindi attivando questa casella. Quest'altra casella, invece, ci mette in grado di dire al nostro cliente di posta elettronica di scaricare la posta in arrivo relativamente a questo *account*, semplicemente cliccando su *Get Mail*. Quindi, quando cliccheremo su *Get Mail*, il nostro *mailer* andrà a scaricare la posta elettronica relativamente a tutti gli *account* che hanno attivato questa opzione. Questa casella, invece, ci permette di dire al nostro programma di posta elettronica se rimuovere oppure no, e a che intervalli di tempo, i messaggi che abbiamo scaricato dal *Mail Server*. Tipicamente ogni *Internet Service Provider* fornisce una *mailbox* di una certa dimensione, dove vanno a finire i messaggi a noi indirizzati. Ovviamente, se noi scarichiamo i nostri messaggi di posta sulla macchina locale, possiamo decidere se rimuovere la copia del messaggio dal *server* o mantenerla. Dobbiamo stare attenti a rimuovere copia del messaggio che abbiamo già letto e scaricato sulla macchina locale, altrimenti la nostra *mailbox* sul *server* che ci fornisce il servizio di posta elettronica potrebbe riempirsi e non essere più in grado di ricevere altri messaggi.

## Advanced 2

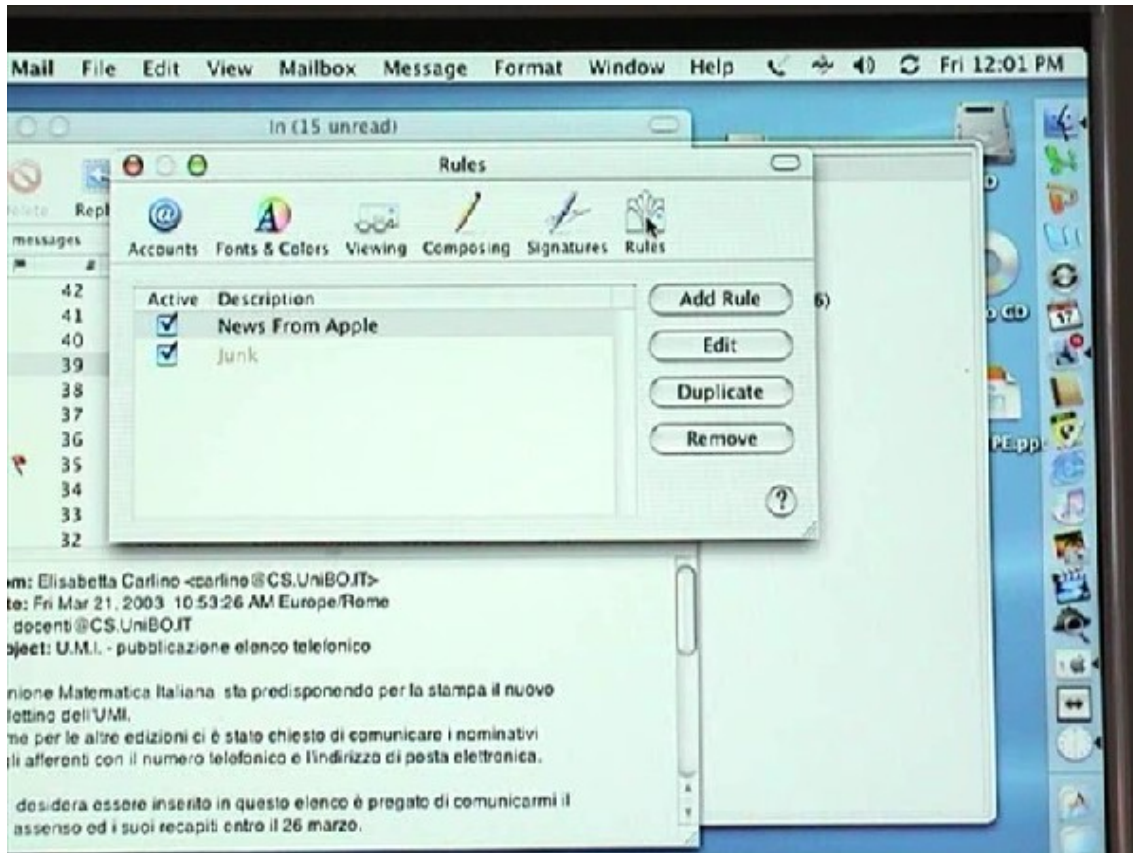




Advanced 2

In questa parte del pannello di configurazione possiamo specificare che grado di sicurezza, nel momento in cui ci autenticiamo presso il nostro *Mail Server*, vogliamo avere. In particolare, possiamo scegliere la porta del *Mail Server* dal quale noi riceviamo la posta elettronica. Su alcune porte del *Mail Server* vengono implementati meccanismi di sicurezza e di autenticazione che assicurano all'utente, ma anche all'organizzazione che fornisce il servizio di posta elettronica, che l'utente che sta cercando di scaricare la posta elettronica sia realmente lui. Quindi, in questo pannello di configurazione, noi sostanzialmente possiamo decidere che grado di sicurezza e di autenticazione vogliamo avere, nel momento in cui andiamo a scaricare la nostra posta elettronica.

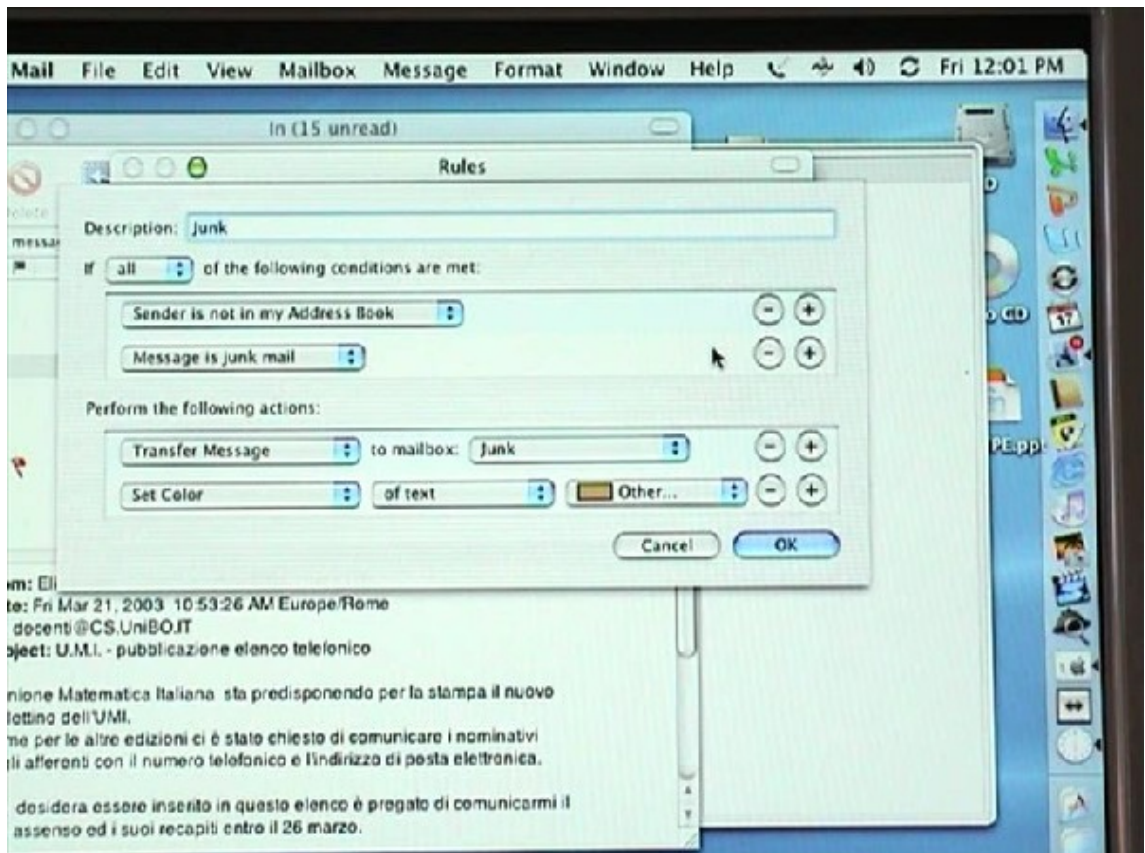
## Rules 1



Rules 1

Mi sembra interessante spendere due parole circa il pannello di configurazione *rules* (regole). In questo pannello possiamo istruire il programma di posta elettronica ad eseguire alcune azioni in corrispondenza di *e-mail* che riceviamo. Ad esempio, la regola di nome *junk* che noi selezioniamo e poi proviamo a modificare, è una regola che ci permette di gestire la posta indesiderata. Questo cosa significa? Significa che il nostro programma di posta elettronica quando riceve un certo *mail*, lo controlla, cerca di accorgersi se questo *mail* è indesiderato e nel caso lo giudichi indesiderato compie un certo numero di azioni.

## Rules 2



Rules 2

Per esempio, nel caso della regola *junk*, se il mittente non è nel mio *address book* e se il messaggio è un messaggio *junk*, in accordo con alcune regola prestabilite del nostro *mailer*, allora trasferisco il messaggio nella *mailbox junk* e lo coloro con questo colore marroncino. Ogniqualvolta noi riceviamo dei messaggi questa regola è attivata, quindi il nostro *mailer*, automaticamente, cerca di capire se questo *mail* è desiderato o no. Se questo *mail* viene considerato come indesiderato, viene trasferito nel *mailbox junk* e viene colorato di marrone. Noi possiamo aggiungere un'altra qualsiasi regola a nostro piacimento e chiamarla con un certo nome, attivarla o disattivarla. Per esempio, potremmo catalogare tutte le *mail* che hanno un certo mittente, trasferendole tutte in una determinata *mailbox*. Per esempio, uno potrebbe decidere di trasferire tutte le *mail* che provengono dal datore di lavoro nella *mailbox* di nome lavoro. La gestione delle *rules* (regole), per uno che fa un uso intenso della posta elettronica, è di fondamentale importanza, perché permette di catalogare automaticamente le *mail* che arrivano classificandole in diverse *mailbox*, oppure cancellandole direttamente. Un altro utilizzo di queste regole potrebbe essere quello di rispondere automaticamente ad un *mail*: se un *mail* proviene da un certo destinatario, rispondi con questo *mail* che io specifico.

## Bibliografia

### Introduzione

#### Protocolli e usare Internet

M.Calvo F.Ciotti G.Roncaglia M.Zela *Internet 2000*; 1999 Laterza

M.Calvo F.Ciotti G.Roncaglia M.Zela *Frontiere di rete*; 2001 Laterza

M.Calvo F.Ciotti G.Roncaglia M.Zela *Internet 2003 Beta*; 2003 Laterza

Sawyer e Williams *Using Information Technology (quarta edizione)*; 2001 McGraw Hill

T. Berners Lee *L'architettura del nuovo Web*; 2001 Feltrinelli

J. Kurose e K. Ross *Internet e Reti di Calcolatori*; 2001 McGraw Hill

W. Lehnert *Internet 101*; 1998 Addison Wesley

P. Lynch e S. Horton *Web Style guide: Basic design principles for creating web sites*; 1999 Yale University Press

Metitieri F. e Manera G. *Dalla Email al Chat Multimediale*; 2000 FrancoAngeli

#### Installazione e configurazione di applicazioni

Giorgio Sitta *Internet Explorer 6*; 2001 Jackson Libri

Doug Lowe *Internet Explorer 6 For Dummies*; 2002 Apogeo - ISBN: 88-7303-978-2

R. Viscardi *Internet Explorer 6 Flash*; 2002 Apogeo - ISBN: 88-503-2012-4

Microsoft Corporation *Microsoft Internet Explorer 6 Resource Kit*; 2002 Microsoft Press - ISBN 88-8331-321-6

Microsoft Italia; <http://www.microsoft.com/italy>

## Approfondimenti

### Motori di Ricerca

Marangoni Roberto e Cucca Alessandro *Motori di ricerca. La via facile e veloce alla ricerca sul Web*; 2001 HOEPLI informatica

Sitta Giorgio *Ricerche su Internet*; 2001 Jackson Libri

Bianchini Massimiliano e Galimberti Andrea *Ricerche on line*; 2000 Alpha Test



Parrini Claudio / *motori di ricerca nel caos della rete. Kit di sopravvivenza tecnico-esistenziale; 2001 Shake*

*Motori di ricerca*; <http://www.motoridiricerca.it>

*The big hub*; <http://www.thebighub.com>

*Cerca*; <http://www.cerca.com>

*Dog pile*; <http://www.dogpile.com>

*Highway 61*; <http://www.highway61.com>

*Metacrawler*; <http://www.metacrawler.com>

## Glossario

**Browser** : programma *client* per navigare sul *Web*.

**Categorie** : una categoria è un insieme di siti *Internet* che trattano uno stesso argomento. Esempi di categoria: sport, politica, informatica, *shopping*, eccetera.

**Chat** : sistema grazie al quale più persone possono conversare in tempo reale in modo sincrono su *Internet*.

**Cliente/client** : programma (o macchina) che richiede servizi a un *server*.

**Cookie** : informazioni che il sito su cui stiamo navigando cerca di memorizzare sull'*hard disk* del nostro *computer*.

**Directory** : strumento di ricerca per siti *Internet* interrogabile utilizzando il sistema delle categorie e sotto-categorie.

**Download** : operazione di trasferimento di informazioni da una macchina remota alla macchina locale.

**File temporanei** : *file* che il *browser* memorizza sull'*hard disk* per un certo lasso di tempo.

**Ftp** : (*File Transfer Protocol*) protocollo di livello applicazione in *TCP/IP* che regola il trasferimento di *file*.

**Gopher** : sistema *client/server* per pubblicare informazioni su *Internet* in formato testuale.

**Helper** : applicazione esterna invocata da un *browser* quando deve gestire un *file* in un formato a lui sconosciuto.

**Help on-line** : istruzioni (suggerimenti, aiuti) fornite da una applicazione (esempio IE) circa il proprio utilizzo. Il manuale di istruzioni è in formato cartaceo. L'*help on-line* invece è in formato digitale ed è consultabile all'interno dell'applicazione che lo fornisce.

**Http** : (*HyperText Transfer Protocol*) protocollo di livello applicazione in *TCP/IP* che gestisce il trasferimento di *file* in formato html su **Internet**.

**Imap** : (*Internet Message Access Protocol*) protocollo che specifica come viene gestita la posta dal vostro *mail server*.

**Internet** : gruppo di reti interconnesse mediante la *suite* di protocolli *TCP/IP*.

**Ipertesto** : sistema di organizzazione delle informazioni (testuali, ma non solo) in una struttura non sequenziale, bensì reticolare.

**Installazione personalizzata** : in una installazione personalizzata l'utente può selezionare a suo piacimento quali moduli installare e quali no. Molte decisioni che in una installazione tipica vengono prese dal programma, in una installazione personalizzata vengono prese dall'utente.

**Installazione tipica** : installazione che risponde alle esigenze standard dell'utente medio. In una installazione tipica il numero di scelte che deve fare l'utente è reso minimo.

**Meta-motori** : siti *Internet* che effettuano ricerche interrogando un certo numero di altri motori di ricerca.

**Motore di ricerca** : strumento di ricerca di risorse *Internet* (non solo siti) interrogabile usando parole chiave.  
Sito che permette di ricercare url di risorse di rete specificando alcune parole chiave.

**Multimedia** : con l'aggettivo multimediale ci si riferisce alla possibilità di inserire contemporaneamente in uno stesso documento informazione di tipo diverso. Ad esempio, immagini, suoni e filmati.

**Operatori** : danno la possibilità di effettuare ricerche su un motore di ricerca combinando in modo opportuno diverse parole chiave.

**Pagina iniziale** : è la pagina *Internet* che verrà scaricata e visualizzata ogni qual volta attiveremo IE sul nostro calcolatore oppure cliccheremo sulla barra dei pulsanti in corrispondenza dell'icona a forma di casetta.

**Parole chiave** : sono parole o frasi utilizzate per ricercare siti *Internet* o più in generale risorse.

**Pop** : (*Post Office Protocol*) protocollo di livello applicazione di *TCP/IP* che gestisce il trasferimento della posta elettronica dal *mail server* alla macchina locale dell'utente.

**Posta elettronica** : sistema asincrono che permette di spedire messaggi testuali (e *file* di ogni tipo) da un mittente a uno o più destinatari di cui si conosce l'indirizzo.

**Plug-in** : applicazione integrata (interna) invocata da un *browser* quando deve gestire un *file* in un formato a lui sconosciuto.

**Protocollo/i** : regole e formati per la comunicazione fra due *computer* in una rete.

**Servente/server** : programma (o macchina) che fornisce servizi ai *client*.

**Smtp** : (*Simple Mail Transfer Protocol*) protocollo di livello applicazione di *TCP/IP* che gestisce il trasferimento di messaggi di posta elettronica su **Internet**.

**Telnet** : protocollo di livello applicazione di *TCP/IP* che permette di effettuare il *login* su una macchina remota.

**Upload** : operazione di trasferimento di informazioni dalla macchina locale ad una macchina remota.

**Url** : (*Uniform Resource Locator*) schema di denominazione delle risorse in **Internet**.

**Web** : sistema informativo ipertestuale e multimediale distribuito su **Internet**.

## Autori

Hanno realizzato il materiale di questo modulo:

### **Prof. Luciano Margara**

Luciano Margara (Palermo 1968) ha conseguito il dottorato di ricerca in Informatica presso l'Università di Pisa, ed è oggi professore associato di Informatica presso l'università di Bologna. Ha svolto attività didattica presso diverse università italiane ed estere. È stato *visiting professor* alla *Berkeley University* e alla *Cornell University*. I suoi contributi di ricerca riguardano la teoria degli algoritmi, i sistemi dinamici complessi a tempo discreto e le euristiche per risolvere in modo approssimato problemi computazionali difficili.