

Appunti Internet

Loris Tissino
© Loris Tissino, 1997–2003

Indice

1 Introduzione	1
1.1 Questi appunti	1
1.2 Licenza	1
2 Introduzione ad Internet	2
2.1 Cos'è Internet	2
2.2 I protocolli	3
2.2.1 Protocolli a livello applicativo	3
2.2.2 Protocolli a livello di trasporto	4
2.2.3 Protocolli a livello di rete	4
2.3 Le interfacce	4
2.4 Gli indirizzi IP	4
2.5 Il DNS (Domain Name System)	5
2.6 Il modello client/server	6
2.7 Cosa si può fare su Internet	6
2.8 Il concetto di URI (Uniform Resource Identifier)	6
3 Il world wide web	9
3.1 Cos'è il World Wide Web	9
3.2 HTTP	9
3.3 HTML	10
3.4 I siti web	11
4 Il browser	12
4.1 Le funzioni di un browser	12
4.2 Tipi di browser	12
4.3 Interfaccia di un tipico browser	13
4.4 In linea e fuori-linea	13
4.5 Utilizzare fuori linea una pagina web	14
4.6 Recupero di pagine web	14
4.7 I server proxy	15
5 Ricerche su web	16
5.1 Come trovare la pagina o il documento che ci interessa	16
5.2 I cataloghi	16
5.3 I motori di ricerca	16
5.4 Alcuni indirizzi	17
5.4.1 Motori di ricerca e cataloghi	17
5.4.2 Motori di ricerca specializzati	17
5.4.3 Meta-motori e cataloghi di motori	18
5.5 I portali	18
6 La posta elettronica	19
6.1 Cos'è e come funziona	19
6.2 Programmi per la posta elettronica	19
6.2.1 La spedizione	19
6.2.2 La lettura tramite prelievo dal server	20
6.2.3 La lettura dei messaggi lasciati sul server	20
6.3 Anatomia di un messaggio email	21
6.3.1 Gli allegati	21
7 I programmi per la posta elettronica	22
7.1 Il concetto di Mail User Agent	22
7.2 Funzioni tipiche di un programma di posta	22
7.3 La composizione di un messaggio	22

Indice

<u>7 I programmi per la posta elettronica</u>	
<u>7.3.1 I destinatari</u>	22
<u>7.3.2 La netiquette</u>	23
<u>7.4 Funzionalità avanzate di un programma di posta elettronica</u>	23
<u>7.4.1 Aspetti organizzativi</u>	23
<u>7.4.2 Account</u>	23
<u>7.5 Aspetti relativi alla sicurezza</u>	24
<u>7.6 Servizi utilizzabili tramite posta elettronica</u>	24
<u>8 Domande ed esercizi</u>	25
<u>8.1 Domande</u>	25
<u>8.2 Esercizi ed attività pratiche</u>	25
<u>9 Licenza d'uso per questo documento</u>	26
<u>9.1 GNU Free Documentation License</u>	26

1 Introduzione

1.1 Questi appunti

Ho iniziato la stesura di alcuni appunti su Internet e l'uso di diverse applicazioni di rete (browser web, programmi di posta elettronica, ecc.) diversi anni fa, quando ho cominciato a tenere corsi di aggiornamento e di formazione professionale sull'argomento. I miei appunti erano molto scarni e avevano esclusivamente lo scopo di fissare alcuni termini.

Qualche mese fa (autunno 2002), ho deciso di dare il mio modesto contributo personale al progetto ECDL-Libre [<http://ecdllibre.sourceforge.net>] iniziando la loro trasformazione in quello che dovrebbe essere un documento collettivo per la preparazione agli esami ECDL sostenuti usando solo strumenti software liberi.

Diverse cose scritte in questo documento vanno oltre le richieste necessarie per il superamento dell'esame relativo al modulo 7 dell'ECDL. Sono però, a mio avviso, essenziali per capire meglio le cose e padroneggiare gli strumenti che si utilizzano.

Quella che pubblico ora è solo una prima bozza di documento, senza immagini e senza esempi espliciti d'uso di qualche browser e programma di posta elettronica. Ovviamente sono in attesa di commenti, considerazioni, suggerimenti, ecc.

Gli aggiornamenti saranno a disposizione in futuro all'indirizzo www.tissino.it/docs/appuntiinternet.

Buona lettura.

L'autore

P.S. Il progetto ECDL-Libre si sta orientando a LaTeX come strumento per la predisposizione dei documenti. Al momento invece questo documento è stato realizzato attraverso script PHP su file HTML che vengono postprocessati con HTMLDOC [<http://www.htmldoc.org>]. Appena possibile convertirò il tutto in formato LaTeX.

1.2 Licenza

```
Copyright (c) 1997-2003 Loris Tissino
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.2
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

2 Introduzione ad Internet

2.1 Cos'è Internet

È difficile dare una definizione corretta di **Internet**, una rete di reti di calcolatori, distribuita su tutti i continenti del pianeta, che comunicano attraverso una serie di protocolli standard tra cui i più importanti sono TCP e IP.

Innanzitutto, si tratta di una **rete di reti**. Quando alcuni calcolatori vengono posti nelle condizioni di comunicare tra di loro, si ha una rete. Una rete può però essere collegata ad un'altra rete. Sostanzialmente, ciò si ottiene facendo in modo che un calcolatore appartenga ad entrambe le reti, e si incarichi di inoltrare il traffico da una rete all'altra. La comunicazione tra due reti può anche essere definita *comunicazione inter-rete*.

A proposito, il calcolatore che appartiene a entrambe le reti si chiama non a caso **gateway**, ossia punto di passaggio. Non ce lo si deve immaginare per forza di cose come un calcolatore tipo quelli con cui normalmente abbiamo a che fare (anche se potrebbe tranquillamente esserlo). L'importante è la sua funzione, non il suo aspetto.

In inglese, si usano due termini diversi per distinguere una qualsiasi **internet** (inter-rete), da **the Internet**, con articolo determinativo e maiuscola: quest'ultima è quella che in italiano chiamiamo normalmente e semplicemente *internet*.

Il fatto che Internet sia diffusa su tutti i continenti è più un dato di fatto che una parte della definizione: Internet era Internet anche prima di essere utilizzata in tutti i continenti e continua ad essere Internet ora che è estesa fino alle stazioni spaziali in orbita.

Una definizione migliore potrebbe avere a che fare con l'evoluzione storica delle reti di calcolatori e delle reti di reti. Semplicemente, dalla fine degli anni '60, quando i sistemisti dei laboratori di ricerca di università e aziende si sono posti il problema di far comunicare tra loro calcolatori, non hanno reinventato la ruota ogni volta, ma hanno sfruttato idee, programmi, metodologie che qualche loro collega aveva pensato prima di loro, magari migliorandole un po' e presentando proposte al vaglio della comunità (i cosiddetti documenti **RFC**, *Request For Comments*).

Quindi, cos'è Internet? È la rete di reti che si è evoluta in questo modo, con una serie di regole condivise e via via migliorate, con nuove reti aggiunte un po' alla volta, fino a diventare *la* rete di reti per antonomasia.

Ci sono un sacco di altre reti di reti che coprono l'intero pianeta ma che non hanno relazioni (quanto meno dirette) con Internet: si pensi ad esempio alla reti per la gestione dei pagamenti con carta di credito, alle reti militari, e così via.

È però importante definire cosa **non è** Internet:

- non è un particolare programma
- non è qualcosa che qualcuno può possedere
- non è un modo per rappresentare le informazioni

Bisognerebbe cercare di evitare, quindi, espressioni del tipo:

"Hai installato Internet sul tuo calcolatore?"

(al massimo, si potrà installare un programma o configurare il calcolatore per *accedere ad Internet*.)

"Hai Internet?"

(non si può *avere* Internet.)

"*Salvami questo documento in formato Internet*"

(non esiste un formato Internet, esistono diversi formati per le diverse cose che si possono fare su Internet.)

2.2 I protocolli

Avendo a che fare con Internet, sentirete spesso parlare di **protocolli**, ossia di regole per la comunicazione tra un calcolatore ed un altro calcolatore o, detto più precisamente, tra un programma (o una componente del sistema operativo) in funzione su un calcolatore ed un corrispondente programma (o componente del sistema operativo) in funzione su un altro calcolatore.

I protocolli servono a definire come questi programmi devono comunicare tra loro: ad esempio, come si inizia la comunicazione, in che formato indicare alcuni dati (ad esempio, la data: anno–mese–giorno o giorno–mese–anno?), cosa si deve fare se la richiesta non può essere accolta, ecc.

Essi descrivono in modo preciso e formale come le comunicazioni devono avvenire. A meno che non si debba scrivere un programma applicativo, non è necessario conoscere i dettagli. Può però rivelarsi molto utile avere almeno una vaga idea di queste cose per poter cercare di definire un problema che si pone, e fare dei test opportuni prima di chiamare un servizio di assistenza tecnica.

I protocolli hanno a che fare con diversi livelli di astrazione.

2.2.1 Protocolli a livello applicativo

Ad un livello alto di astrazione, possiamo immaginare i calcolatori che riescono facilmente a scambiarsi messaggi di testo e documenti vari. È a questo livello, detto **livello applicativo**, che sono definiti i protocolli di comunicazione per le attività più immediatamente riconoscibili dall'utente (accesso a siti web, posta elettronica, ecc.)

Ad esempio, il protocollo **HTTP** prevede che un programma presente su un calcolatore (tipicamente il *browser web*, ne parleremo più avanti) mandi ad un altro programma (il cosiddetto *server web*) una richiesta del tipo [qui la comunicazione è tradotta in italiano ed un po' interpretata]:

```
Sto usando la versione 1.1 del protocollo HTTP
Mandami il documento test.html che hai a disposizione nella cartella documenti
Il mio nome è Librowser 2.2
Mi sto rivolgendo a te, server web che rispondi al nome www.esempio.it
Mandami il documento solo se è stato modificato dopo il 2 gennaio 2003
```

Proseguendo nell'esempio, il programma server potrebbe rispondere:

```
Sono il server web FreeHTTP Daemon
Questo documento non lo trovo
```

oppure:

```
Sono il server web FreeHTTP Daemon
Ecco il documento
L'ultima modifica risale al 14 gennaio 2003
Ti informo che si tratta di una pagina web (un file di testo in formato html)
```

(segue il documento)

2.2.2 Protocolli a livello di trasporto

Ad un livello di astrazione più basso, possiamo voler definire come un intero blocco di dati (un file), anche di grandi dimensioni, debba essere trasferito da un calcolatore ad un altro.

Così come se si deve spostare un armadio da una stanza ad un'altra può essere necessario smontare l'armadio, numerare i pezzi, spostare i pezzi separatamente l'uno dall'altro e rimontare l'armadio a destinazione (controllando che nessun pezzo si sia perduto durante il tragitto, ed eventualmente cercando di recuperarlo), quando si devono trasferire molti dati, questi vengono suddivisi in *pacchetti* che, dopo essere stati numerati, vengono trasferiti indipendentemente e riassemblati una volta arrivati. Il protocollo che descrive questo procedimento si chiama *TCP*, *Transmission Control Protocol*.

2.2.3 Protocolli a livello di rete

Ad un livello di astrazione ancora più basso possiamo voler definire come fare a organizzare le reti di reti di calcolatori in maniera tale da consentire ad un singolo pacchetto di passare dal calcolatore di origine a quello di destinazione. Di ciò si occupa l'*IP*, *Internet Protocol*.

2.3 Le interfacce

Mentre un protocollo definisce come un programma comunica con un altro programma suo pari (il browser web con il server web, il programma di posta con il server di posta, ecc.), un'*interfaccia* descrive come un programma comunica con un altro programma di diverso livello nell'ambito dello stesso calcolatore.

Vi ricordate i livelli di astrazione visti prima? È un po' come se si dicesse che un programma di alto livello non volesse avere a che fare con i dettagli (suddivisione di un file in pacchetti, trasmissione dei pacchetti, ecc.) e delegasse a livelli software inferiori l'esecuzione di questi compiti. La descrizione di questo tipo di comunicazione è appunto l'interfaccia.

La questione potrebbe risultare più chiara pensando ad un livello ancora superiore, quello della comunicazione fra utente (essere umano) e programma applicativo (browser web, programma di posta). L'utente umano non vuole avere a che fare con i dettagli di funzionamento del programma e di come questo compirà le operazioni che vengono richieste, e agirà tramite un'interfaccia (ad esempio, un'interfaccia grafica con icone, pulsanti e menù, oppure un'interfaccia a riga di comando testuale, oppure ancora un'interfaccia vocale).

È importante notare che il tipo di interfaccia utilizzata da una delle parti in comunicazione può essere tranquillamente diversa da quella utilizzata dall'altra parte: ad esempio, io posso usare un programma da riga di comando (interfaccia testuale) per spedire un messaggio di posta elettronica, che potrà essere letto dal destinatario con un programma di posta che fa uso di un'interfaccia grafica, e viceversa.

2.4 Gli indirizzi IP

Uno dei problemi da risolvere per permettere la comunicazione tra calcolatori in una rete di reti quale è Internet è quello di garantire che ciascun calcolatore abbia un identificativo univoco, in modo da poter indicare chiaramente a chi (a quale calcolatore) deve essere indirizzato un determinato pacchetto.

Il metodo scelto per risolvere il problema consiste nell'assegnare a ciascun calcolatore collegato ad Internet un indirizzo univoco, detto *indirizzo IP*, per ogni rete a cui quel calcolatore appartiene [cioè viene assegnato un indirizzo a ciascuna *interfaccia di rete* di quel calcolatore].

Un indirizzo IP è un numero di quattro byte, solitamente indicati nella loro forma decimale: ad esempio, *62.177.1.107*. Il software che si occupa della trasmissione dei pacchetti dal calcolatore ad un altro viene

informato del fatto che il calcolatore appartiene ad una determinata rete (ad esempio, la rete in cui tutti i calcolatori hanno indirizzi che iniziano per `62.177` [cosa che si fa con una *maschera di sottorete*], e che se deve mandare qualche comunicazione a calcolatori di altre reti dovrà rivolgersi al calcolatore che farà le veci di intermediario (*gateway*), ad esempio al calcolatore con indirizzo `62.177.0.1`.

Ovviamente, sarà necessario garantire che non ci siano due calcolatori su Internet con lo stesso indirizzo IP. Come si ottiene questo? Prima di assegnare un indirizzo ai calcolatori della mia rete, devo farmeli indicare da un'apposita autorità (IANA).

Se però devo assegnare degli indirizzi IP ad una rete privata (con calcolatori che non saranno quindi direttamente raggiungibili da Internet), posso utilizzare degli apposite *classi di indirizzi per reti private* (`10.x.y.z`, `172.16.x.y`, `192.168.x.y`).

2.5 Il DNS (Domain Name System)

Visto che ricordarsi come raggiungere un calcolatore tramite il suo indirizzo è un po' scomodo, nel 1984 è stato inventato un sistema che permette di associare un indirizzo mnemonico (facile da ricordare) ad un calcolatore. Questo sistema si basa su *nomi di dominio* e viene spesso indicato con il suo nome in inglese (*Domain Name System*).

Il sistema funziona all'incirca così: supponiamo che io sia l'amministratore di rete di una società che si chiama *Completamente Inventata*. Potrei chiamare il mio calcolatore *completamenteinventata*, ma se lo facessi non avrei grande sicurezza del fatto che qualcun altro, da qualche parte del mondo, non faccia la stessa cosa. Per evitare problemi, devo *registrare* il cosiddetto *nome di dominio* presso una delle autorità competenti.

Se mi rivolgo all'autorità italiana registrerò il nome di dominio *completamenteinventata.it*, rivolgendomi a quella francese potrei ottenere *completamenteinventata.fr*, e via di seguito.

Le autorità statunitensi che si occupano di questo tipo di registrazioni operano di fatto a livello internazionale e registrano nomi di dominio che terminano con i suffissi *.com*, *.net*, *.org*, *.info*, *.biz*, *.name* (e quindi potrei registrare *completamenteinventata.com*).

A proposito, evitiamo di chiamare questi suffissi "estensioni". Per essere precisi, si dovrebbe parlare di domini di livello primario (*top level domains*).

Una volta registrato il nome di dominio (di secondo livello) per i miei calcolatori, potrò assegnare a ciascuno di essi un nome a mia scelta (ad esempio, *primo.completamenteinventata.it*, *secondo.completamenteinventata.it*). Posso anche attribuire allo stesso calcolatore più nomi se lo desidero [tecnicamente, si parla di *alias*].

Una prassi molto diffusa prevede che ad un calcolatore venga assegnato un nome corrispondente al servizio che quel calcolatore offre: il mio server web si chiamerà quindi (anche) *www.completamenteinventata.it*, il mio server di posta *mail.completamenteinventata.it*, e così via. Non si sottolineerà mai abbastanza che si tratta però solo di una prassi. Nulla impone che un server web debba avere un nome che inizia con *www*, né che un calcolatore il cui nome inizia per *www* ospiti un server web.

Spesso i calcolatori collegati ad internet hanno un proprio nome (il cosiddetto *host name*) che, legato al dominio a cui essi appartengono, forma il *nome di dominio pienamente qualificato* (in inglese, *fully qualified domain name*).

Il procedimento per ottenere la *traduzione* di un nome di dominio pienamente qualificato (ossia per sapere a quale indirizzo IP corrisponde un determinato nome di dominio pienamente qualificato) viene chiamato

risoluzione del nome e viene svolto attraverso il colloquio con dei **server DNS** che assolvono proprio a questo compito. [I server DNS lavorano in collaborazione tra loro e sono organizzati in gerarchie].

2.6 Il modello client/server

Nella terminologia delle reti, si fa spesso uso dei termini **client** (cliente) e **server** (servente), che vengono però usati in contesti diversi e con significati, quindi, diversi.

Tento di spiegarmi con un paio di esempi.

Se dico: "*La stampante da usare è quella collegata al server*" è evidente che intendo con la parola *server* un calcolatore, e in particolare quello che all'interno di una rete offre particolari servizi (si parla di *file server* per la memorizzazione di archivi, *print server* per la gestione delle code di stampa, e così via).

Se invece dico: "*Il server web non risponde*" non parlo del calcolatore fisico, ma del programma server (si dice anche demone) in funzione su di un calcolatore.

In pratica, quando si parla di **reti client/server** si intende che alcuni calcolatori (i serventi) svolgono servizi per altri calcolatori (i clienti). Questo fatto viene sottolineato spesso in contrapposizione con le reti paritetiche (*peer-to-peer*), in cui invece tutti i calcolatori sono sullo stesso piano.

Quando, invece, si parla di **applicazioni client/server** si intende il fatto che esistono programmi che comunicano tra loro: il programma cliente chiede un servizio o un'informazione al programma servente, quest'ultimo esegue e informa il cliente dell'esito dell'operazione. Non di rado, programma cliente e programma servente operano sullo stesso calcolatore.

Gran parte delle cose che si possono fare con Internet funzionano secondo il modello delle applicazioni client/server.

2.7 Cosa si può fare su Internet

Contrariamente a quanti molto pensano, Internet non è costituita solo da World Wide Web (il famoso *www*) e posta elettronica. Nella grande rete si può anche:

- controllare un elaboratore remoto (*telnet, ssh*)
- trasferire file (*file transfer protocol*)
- partecipare a conferenze virtuali (*newsgroups e mailing lists*)
- comunicare in tempo reale sia in modalità testo (*chat, IRC*), sia in audio/video
- utilizzare applicazioni di rete (*desktop virtuali*)
- accedere a contenuti multimediali (*audio e video streaming*)
- condividere documenti (*peer-to-peer file sharing*)

...e fare molto altro. Molte applicazioni usano protocolli di comunicazione basati su TCP e IP.

2.8 Il concetto di URI (Uniform Resource Identifier)

Vista la gran varietà di cose che si possono fare su Internet, è nata l'esigenza di uniformare (standardizzandoli) i modi per riferirsi ad una particolare risorsa. Sono nati così gli identificatori uniformi di risorse o **URI**.

Un URI identifica uno schema di comunicazione e un modo per utilizzare la particolare risorsa che si vuole identificare. Tale modo è dipendente dal tipo di risorsa. Ad esempio, nel caso di pagine web verrà identificato il nome del server web e l'indirizzo completo della pagina (in questo caso si parla di **URL**, *Uniform Resource Locator*, perché la risorsa viene *localizzata*, nel senso che si indica *dove* la si trova).

2 Introduzione ad Internet

Alcuni esempi:

`ftp://ftp.nevada.edu/pub/guitar`

(indica un server FTP da dove è possibile prelevare dei file oppure metterceli)

`telnet://argo.temple.edu`

(indica un host che accetta connessioni in emulazione di terminale)

`http://www.emptywebsite.com`

(indica l'indirizzo di un sito web)

`mailto:info@linux.it`

(indica un indirizzo al quale è possibile mandare un messaggio di posta elettronica)

`news:soc.culture.italian`

(indica il nome di un gruppo di discussione al quale è possibile accedere).

Gli URI relativi alle pagine web meritano un approfondimento, visto che sono quelli con cui si ha più a che fare.

`http://www.organizzazionezeta.it/documenti/iniziative.html`

indica che ci si deve rivolgere ad un server web (*http*) ospitato dal calcolatore con nome DNS

www.organizzazionezeta.it per chiedere un documento (presumibilmente una pagina html) di nome *iniziative.html* presente nella cartella *documenti*.

`http://www.organizzazionezeta.it/documenti/elenco.php`

`http://www.organizzazionezeta.it/documenti/elenco.asp`

`http://www.organizzazionezeta.it/cgi-bin/elenco`

sono riferimenti non ad un documento già presente nel *file system* del calcolatore ma ad un programma che deve essere eseguito su di esso; nel primo caso si tratta presumibilmente di un programma *php* (*PHP: Hypertext Preprocessor*, nel secondo caso di un programma *asp* (*Active Server Pages*, nel terzo di un'applicazione scritta in qualche linguaggio come Perl o C). I server web sono configurati in modo da poter riconoscere (in base a dove si trova il file o alla sua estensione) se il file il cui nome è indicato deve essere inviato così com'è o eseguito.

`http://www.organizzazionezeta.it/mostra.php?articolo=18&lingua=it`

indica al programma che deve essere eseguito sul server alcuni parametri, che vengono presi in considerazione al momento dell'esecuzione; nell'esempio, si potrebbe trattare di un programma che mostra l'articolo con codice 18 nella versione in italiano; la parte che segue il punto interrogativo si chiama *stringa di interrogazione* (*query string*).

`http://www.organizzazionezeta.it/statuto.html#probiviri`

indica che deve essere richiesto al server web il documento *statuto.html*, ma che esso dovrà essere mostrato dal browser a partire dal punto indicato come *probiviri*; la parte che segue il segno # viene chiamata *frammento*.

`http://giorgio@www.organizzazionezeta.it/documentozeta/`

indica che il server web deve essere informato che a fare la richiesta è un utente di nome *giorgio*.

`http://giorgio:pulcinella@www.organizzazionezeta.it/documentozeta/`

indica che il server web deve essere informato che a fare la richiesta è un utente di nome *giorgio*, che usa la password *pulcinella* per accedere ai servizi offerti.

`http://www.organizzazionezeta.it:3240`

2 Introduzione ad Internet

indica che la richiesta va fatta ad un server web che non è in ascolto sulla porta standard (80), ma sulla porta 3240 [la *porta* è un modo per indicare, attraverso un numero, a quale applicazione servente in esecuzione su di un calcolatore ci si vuole rivolgere con una richiesta da parte di un programma cliente].

3 Il world wide web

3.1 Cos'è il World Wide Web

Il *World Wide Web* (letteralmente, la ragnatela mondiale [si noti che il termine inglese *worldwide* è una parola unica, ma è stata spezzata per ottenere l'acronimo WWW]) è nato all'inizio degli anni '90 per venire incontro all'esigenza di rendere la documentazione (scientifica, in principio) facilmente accessibile attraverso Internet. I principi fondamentali alla sua base sono:

- **ipertestualità**, ossia la possibilità di consultare un documento diverso seguendo dei collegamenti tra un punto nel documento (questo punto viene a volte definito *punto sensibile*, *parola calda* o, più semplicemente, *link*) che si sta consultando ed un documento ad esso collegato;
- **standardizzazione del formato** per la rappresentazione dei contenuti, ossia presenza di un modo relativamente semplice (da capire e da applicare) per preparare i documenti che si vogliono pubblicare, indipendente dalla piattaforma (ossia utilizzabile con i più svariati tipi di calcolatori e sistemi operativi) e da un particolare programma;
- **predisposizione per la multimedialità**, ossia capacità di trattare sia documenti testuali tradizionali sia qualsiasi altro tipo di contenuto digitale (immagini, audio, video); i programmi possono anche scambiarsi informazioni a vicenda sul tipo preferito (nella richiesta) e sul tipo mandato (nella risposta), in modo da rendere più facile la vita dell'utente, che non debba preoccuparsi troppo del tipo di contenuto che sta ricevendo.

Oggi questi aspetti ci potrebbero apparire del tutto scontati e naturali, per cui vale la pena di riflettere su come stavano le cose prima della nascita del www.

- benché il concetto di ipertesto fosse già ben noto, i documenti disponibili in rete non erano ipertestuali; leggendo un documento potevo trovare il riferimento ad una legge o ad un trattato, ma non potevo seguire, con la facilità odierna, un collegamento ipertestuale per leggere quella legge o quel trattato: dovevo recuperarlo attraverso una serie di attività (ovvero impartendo una serie di comandi), ammesso di trovare delle indicazioni sul dove trovarlo;
- non esisteva nessuno standard di fatto su come presentare documenti agli altri: chi usava un programma di elaborazione dei testi presentava i documenti semplicemente nel formato tipico di quel programma, e il documento poteva essere letto solo da chi possedeva un programma in grado di capirlo [a volte questo capita ancora, ma fino a metà degli anni '90 succedeva sempre]; l'unica alternativa era il formato cosiddetto *plain text*, testo semplice in cui si rinuncia(va) ad ogni tipo di formattazione;
- le immagini a corredo dei testi (per non parlare di altri contenuti) erano sempre separate da essi; per qualsiasi file ricevuto si doveva cercare di capire con quale programma esso poteva essere trattato.

Cosa c'entra una **ragnatela** in tutto questo? Semplicemente, il termine *rete* rappresenta il fatto che i calcolatori sono collegati tra di loro e possono comunicare, mentre con **ragnatela** si voleva sottolineare l'esistenza di un intreccio (molto più fitto) di collegamenti ipertestuali fra un documento ed un altro.

Per ottenere quello che conosciamo come www o, più semplicemente, **web**, sono stati studiati un protocollo di comunicazione (HTTP) ed un linguaggio per la descrizione dei contenuti (HTML).

3.2 HTTP

Il protocollo **HTTP** (HyperText Transfer Protocol) prevede che un programma cliente si rivolga ad un programma server per chiedere un documento che il server mette a disposizione, oppure la generazione di un documento in base a particolari parametri forniti.

3 Il world wide web

Il programma cliente tipico, utilizzato dagli utenti umani, è detto **browser** (letteralmente, *sfogliatore*). Il documento tipico ricevuto è detto **pagina web**.

Il programma server (normalmente chiamato **web server** o **demone http**) invia una risposta con un'intestazione (che tipo di documento sto mandando, quanto è grande, ecc.) e poi il documento effettivo.

A volte il server web non ha a disposizione un documento già bell'e pronto. In alcuni casi (ad esempio per un errore nella richiesta) risponderà che il documento non esiste, o si trova ad un altro indirizzo; in altri genererà un documento "al volo" (ad esempio, in funzione della lingua preferita indicata nella richiesta).

Non sempre le richieste arrivano da browsers; a volte ad effettuarle sono programmi di scansione automatica dei contenuti (i cosiddetti **bots**, **spiders** o **crawlers**).

3.3 HTML

Le pagine web sono scritte in un linguaggio chiamato **HTML** (*HyperText Markup Language*).

Ad esempio, un semplice documento HTML di questo tipo:

```
<HTML>
<HEAD>
<TITLE>Appunti Internet</TITLE>
<META HTTP-EQUIV="Content-Type" CONTENT="text/html; CHARSET=iso-8859-1">
</HEAD>
<BODY>
<DIV ALIGN="center">
<P><IMG SRC="pc.gif" ALT="Immagine di un PC" HEIGHT=110 WIDTH=110></P>
<P><STRONG>Questa è una pagina Web</STRONG></P>
<P>Questo è un <A HREF="http://www.emptywebsite.com">link
ad una pagina presso un diverso server</A></P>
</DIV>
</BODY>
</HTML>
```

verrebbe rappresentato in questo modo da un browser:

3.4 I siti web

Per *sito web* si intende un insieme di pagine web curate da una organizzazione in modo unitario e coordinato.

Il sito web non va confuso con:

- una singola pagina web
- un nome di dominio (un sito può essere costituito da più domini, con un dominio si possono avere più siti, per cui i termini non sono sovrapponibili)
- una casella di posta elettronica

Andranno quindi evitate espressioni tipo:

- *"Ho visto il sito web all'indirizzo <http://www.qualsiasiosa.inv/paginaparticolare.html>"* (si fa riferimento non al sito intero ma ad una sua componente);
- *"Registriamo il sito per la nostra azienda"* (si registra un dominio, non un sito);
- *"Mandami un messaggio email al mio sito mario@postavaria.inv"* (c'è bisogno di spiegazioni?)

4 Il browser

4.1 Le funzioni di un browser

Il *browser web* è il programma che l'utente usa per accedere a documenti presenti nel world wide web. Il termine *browser* letteralmente significa "sfogliatore" (visto che si usa per sfogliare le pagine web). In italiano a volte si usa il termine "navigatore" e si dice "navigare nel web".

Un *browser* deve essere in grado di:

- inviare richieste ad un web server utilizzando il protocollo HTTP;
- ricevere e interpretare correttamente la risposta (intestazione e documento ricevuto);
- se il documento ricevuto è rappresentabile direttamente, mostrarlo all'utente (nel caso di pagine web, interpretando il codice HTML in cui sono scritte e formattandole in maniera adeguata);
- se il documento ricevuto è una pagina web che contiene riferimenti ad immagini o altri file da visualizzare all'interno della pagina, richiedere anche questi;
- gestire i collegamenti ipertestuali, segnalando in maniera opportuna la presenza degli stessi e permettendo all'utente di seguirli (molti browser grafici lo fanno, in assenza di diverse indicazioni, con un puntatore del mouse che assume la forma di una mano);
- permettere la compilazione di moduli da parte dell'utente (per modulo si intende lo spazio all'interno di una pagina in cui esistono dei campi dove l'utente può inserire delle informazioni);

Altre funzioni che normalmente vengono messe a disposizione da un browser sono le seguenti:

- gestione di una *cache* dei documenti ricevuti (in modo da poterli riutilizzare senza richiederli di nuovo al server oppure controllando semplicemente che nel frattempo non siano stati modificati);
- gestione di una rubrica di indirizzi in modo che sia facile per l'utente tornare a visitare una pagina (se ha aggiunto l'URL di una pagina alla rubrica); tale rubrica viene normalmente indicata come raccolta di *segnalibri*, ma un noto *browser*, in maniera non corretta, la chiama raccolta di *preferiti* [potrei volermi ricordare l'indirizzo di una pagina web proprio perché non mi piace: che senso ha dire che è una mia "preferita"?];
- possibilità di stampare un documento (completamente o parzialmente);
- possibilità di salvare un documento su disco;

Inoltre, i browser devono essere in grado di gestire un certo livello di configurazione da parte dell'utente (indicazione della lingua preferita, della dimensione dei caratteri, ecc.), alcuni aspetti relativi alla sicurezza (gestione dei contenuti "attivi", cioè di programmi di cui viene richiesta l'esecuzione al calcolatore che li riceve), alla gestione delle sessioni di lavoro, ecc.

4.2 Tipi di browser

Esistono molti browser, che si differenziano sia dal punto di vista tecnico (cosa fanno, come lo fanno) sia dal punto di vista della distribuzione e della politica di licenze (come per tutti i tipi di programmi).

I browser si differenziano per l'interfaccia (grafica o testuale), la configurabilità (più o meno spinta), il supporto per gli standard, e così via. Come per tutte le categorie di applicazioni, è più importante capire le funzionalità generali che imparare scorciatoie e pulsanti. In questo documento faremo riferimento ad un browser libero, *Mozilla*, ma i concetti appresi potranno essere facilmente adattati ad altri programmi della categoria.

4.3 Interfaccia di un tipico browser

Le funzioni di base di un browser sono abbastanza standardizzate. Tramite una "barra degli strumenti", i menù, particolari "scorciatoie da tastiera", o interagendo con la pagina visualizzata, è normalmente possibile:

- scegliere di lavorare *on-line* (in linea) oppure *off-line* (fuori linea), concetti che approfondiremo successivamente;
- impostare l'indirizzo di una pagina web che si vuole vedere;
- seguire i collegamenti ipertestuali;
- tornare alla pagina precedente;
- tornare alla pagina successiva;
- arrestare il caricamento di una pagina;
- aggiornare una pagina (chiedendo al browser di verificare se la pagina è stata modificata sul server, facendosela mandare in caso affermativo);
- visualizzare e gestire riquadri (*frames*), che dividono lo spazio nella finestra del browser tra più pagine web;
- andare ad una pagina impostata come "predefinita";
- impostare una determinata pagina come "predefinita";
- segnare l'indirizzo della pagina corrente in un proprio elenco di *segnalibri*;
- visualizzare un elenco delle pagine visitate recentemente;
- accedere facilmente ad un sito web che permette di effettuare delle ricerche [normalmente è possibile scegliere qual è il sito web da richiamare]
- registrare su disco una pagina;
- stampare una pagina o una sua parte;
- visualizzare il codice sorgente (HTML) della pagina;
- compilare moduli (*forms*);
- personalizzare i livelli di sicurezza e le impostazioni di visualizzazione delle pagine (ad esempio, decidendo di mostrare o meno le immagini);

4.4 In linea e fuori-linea

Essere in linea di per sè significa essere collegati ad Internet (ad esempio, con il modem acceso ed un collegamento al proprio Internet Service Provider, oppure con un'interfaccia di rete attivata). Mettere fuori linea il proprio browser non vuol dire però scollegarsi da Internet ("staccare" il modem), bensì indicargli di non rivolgersi a server web per richiedere documenti, e di utilizzare solo documenti presenti nel file system locale (normalmente, quelli ottenuti durante precedenti attività, o quelli di un cd-rom).

In pratica, sono possibili queste quattro combinazioni:

	browser in linea	browser non in linea
connesso ad Internet	i link funzionano, si possono vedere tutte le pagine del WWW, ecc.	si possono vedere solo le pagine presenti nella cache, ma si continua a sostenere il costo della connessione (però forse abbiamo altri programmi in esecuzione che la richiedono, come ad esempio un client e-mail o FTP)
sconnesso da Internet	si possono seguire solo i collegamenti ipertestuali che portano a pagine presenti nel file system locale; alcuni browser segnalano l'impossibilità di seguire gli altri e propongono il passaggio alla modalità in linea se si tenta di seguire gli altri collegamenti ipertestuali	si possono vedere solo le pagine presenti nel file system locale o nella cache, ma senza sostenere il costo della connessione (ottima modalità di lavoro quando ad esempio si consultano pagine web presenti nel proprio hard disk o su di un cd-rom)

4.5 Utilizzare fuori linea una pagina web

Molto spesso si accede al WWW per cercare documenti su qualche argomento di interesse. Una volta trovato il documento, per poterlo utilizzare (leggere, stampare, importare, ecc.) sono possibili diverse strade:

- semplice ***copia & incolla*** di parte del testo (tra il browser ed un'altra applicazione, ad esempio un elaboratore di testi)
- registrazione del documento in ***formato solo testo*** (*plain text*), con la perdita di formattazioni, link, ecc.
- registrazione del documento in ***formato html***, così come l'ha ricevuto il browser
- registrazione del documento come ***pagina web completa***, con il mantenimento di formattazioni e immagini in file separati
- ***importazione*** del sito web, appoggiandosi a programmi esterni specializzati

Alcuni browser mettono a disposizione anche modalità diverse, come il salvataggio in formato posta elettronica e la possibilità di predisporre operazioni di sincronizzazione (a intervalli regolari o su richiesta il browser controlla che i file nella propria cache corrispondano agli originali).

4.6 Recupero di pagine web

A volte può tornare utile cercare di riprodurre siti web esterni sul proprio hard disk. L'operazione non sempre è fattibile per motivi tecnici, ma è bene sapere che esistono dei programmi che si possono tentare di sfruttare. Alcuni funzionano cercando di ricreare la struttura della directory esterna in locale "copiando" i file. Altri si appoggiano ad un browser "comandandogli" il recupero di pagine web collegate a quelle di partenza. Altri ancora funzionano in maniera analoga ad un proxy server.

Se si dispone di un elenco di URL di pagine web che si desiderano leggere, a volte potrebbe essere conveniente sfruttare programmi come ***wget*** dalla riga di comando o tramite un front-end grafico.

Ad esempio, per fare una copia dell'intero contenuto di un sito web si può utilizzare il comando

```
wget --mirror --no-parent http://www.qualchesito.xy
```

[wget supporta molte opzioni per la personalizzazione, si consulti la guida in linea per i dettagli]

4.7 I server proxy

In alcune configurazioni di rete, viene suggerito di utilizzare una connessione tramite *server proxy*. Il termine inglese *proxy* significa "procuratore, intermediario". Sostanzialmente, un server proxy è un programma che lavora come intermediario tra il nostro browser e il server web remoto. Il browser non invia direttamente la richiesta al server web remoto, ma la invia al server proxy, il quale la inoltra a destinazione. Quando arriva la risposta, questa viene fornita al browser ha fatto la richiesta iniziale.

Un'inutile complicazione? No, perché i server proxy possono svolgere alcuni utili compiti:

- possono gestire una cache collettiva delle pagine web visitate da tutti gli utenti di una rete
- possono applicare dei filtri basati sugli indirizzi (URI) o sui tipi di contenuti
- possono mantenere un registro delle risorse richieste e delle attività svolte

5 Ricerche su web

5.1 Come trovare la pagina o il documento che ci interessa

Come arrivare, nell'enorme "mare" del WWW, al link che ci porterà al documento che stiamo cercando? Probabilmente lo troveremo in:

- Media tradizionali (riviste, giornali, tv)
- Messaggi di posta elettronica
- Pagine di altri siti
- Cataloghi
- Motori di ricerca
- Motori di ricerca specializzati per argomento
- Meta-motori

5.2 I cataloghi

Per *catalogo* si intende un elenco ragionato di risorse presenti sul web, suddivisi in categorie e commentati (in genere con un lavoro redazionale compiuto da esseri umani). In inglese si usa il termine *directory*.

Un esempio: [Dmoz.org](http://www.dmoz.org) [<http://www.dmoz.org>] [immagine]

Può essere molto utile consultare i cataloghi quando si ha idea di cosa si sta cercando ma non si sa bene che nome abbia. Nel catalogo, ci si può avvicinar per approssimazioni successive a ciò che si sta cercando.

Ad esempio, ho sentito parlare di un nuovo linguaggio di programmazione, ma non mi ricordo il suo nome. In un catalogo, posso cercare nella voce "Computer", poi in quella "Software", infine "Linguaggi di programmazione". Se sono fortunato, vedrò apparire nell'elenco il nome che stavo cercando.

5.3 I motori di ricerca

I *motori di ricerca* sono siti web che permettono di fare delle ricerche in enormi indici creati ed aggiornati nel tempo mediante la scansione, da parte di agenti software appositamente creati, di un numero considerevole di pagine web.

Quando cerco qualcosa utilizzando un *motore*, cerco informazioni su un indice presso il motore, non direttamente nelle pagine web che potrebbero avere il contenuto che sto cercando. Questo fatto è da tenere presente perché ci potrebbero essere risultati non validi a causa di aggiornamenti delle pagine web.

Di solito i motori di ricerca si presentano con un'interfaccia semplice in cui è presente un campo di input in cui l'utente digita una *espressione di ricerca*, eventualmente accompagnata da delle opzioni, e da un pulsante con la scritta "Cerca".

Inserendo l'espressione e cliccando sul pulsante ci verrà presentata una lista di indirizzi che potrebbero fare al caso nostro, accompagnati da informazioni aggiuntive che ci dovrebbero permettere di capire se siamo vicini o lontani dalla soluzione desiderata (a leggere bene queste informazioni, è un po' come se ci dicessero 'acqua', 'fuochino', 'fuoco!...').

In che ordine è presentata questa lista? Non c'è una ricetta standard: ogni motore di ricerca attribuisce dei punteggi alle pagine elencate in base ad alcuni criteri considerati con diverse priorità. Ad esempio, potrebbero far aumentare il punteggio di una pagina i seguenti fattori:

5 Ricerche su web

- le parole che sto cercando compaiono nel titolo
- le parole che sto cercando compaiono più volte nel testo
- le parole che sto cercando sono tra loro vicine
- le parole che sto cercando sono nella parte superiore della pagina
- le parole che sto cercando costituiscono un collegamento ipertestuale
- la pagina in considerazione è riferita (ossia vi si giunge seguendo collegamenti ipertestuali) da molte altre pagine nel web

Sostanzialmente, il trucco per fare delle ricerche consiste nel provare a fare delle ipotesi su ciò che potrebbe esserci scritto in un documento che parla della cosa che sto cercando, allargando o restringendo i criteri di ricerca a seconda dei risultati ottenuti.

Si tenga presente che i motori di ricerca presentano spesso un'interfaccia di ricerca avanzata in cui si possono precisare:

- i legami fra i termini che sto cercando (a volte tramite gli operatori logici AND e OR, a volte utilizzando altre sintassi)
- la lingua in cui il documento che sto cercando dovrebbe essere scritto (cosa molto utile se sto cercando in base al nome di qualcuno o in base ad un termine usato in più lingue)
- altre cose che dipendono dal singolo motore

Ogni motore di ricerca mette a disposizione solitamente una *guida all'uso* che propone una serie di esempi al riguardo. Raccomando la lettura di alcune di queste.

5.4 Alcuni indirizzi

5.4.1 Motori di ricerca e cataloghi

- Google [<http://www.google.it>]
- Yahoo! [<http://www.yahoo.it>]
- Lycos [<http://www.lycos.it>]
- Altavista [<http://www.altavista.it>]
- Hotbot [<http://www.hotbot.com>]
- Northernlight [<http://www.northernlight.com>]
- Arianna [<http://www.arianna.it>]
- Virgilio [<http://www.virgilio.it>]
- Infoseek [<http://www.infoseek.com>]
- Excite [<http://www.excite.com>]
- Euroseek [<http://www.euroseek.com>]

5.4.2 Motori di ricerca specializzati

Ovviamente esistono "motori di ricerca" specializzati per argomento. Alcuni esempi:

- download.com [<http://www.download.com>] (software)
- WinFiles [<http://www.winfiles.com>] (applicazioni e suggerimenti per Windows 95)
- IconBazaar [<http://www.iconbazaar.com>] (icone)
- Deja [<http://www.deja.com/news/>](newsgroups)
- Liszt [<http://www.liszt.com>] (mailing lists)
- FAQs.org [<http://www.faqs.org>] (frequently asked questions)

5.4.3 Meta–motori e cataloghi di motori

I meta–motori permettono più ricerche contemporanee:

- [Ask Jeeves](http://www.ask.com) [*http://www.ask.com*]
- [MetaCrawler](http://www.go2net.com/search.html) [*http://www.go2net.com/search.html*]
- [SavySearch](http://www.savvysearch.com) [*http://www.savvysearch.com*]
- [Copernic](http://www.copernic.com) [*http://www.copernic.com*]

Una guida ai motori di ricerca e suggerimenti sul loro uso:

- [MotoriDiRicerca.it](http://www.motoridiricerca.it) [*http://www.motoridiricerca.it*]

5.5 I portali

I *portali* sono siti Internet che si propongono come "punto d'accesso" per la navigazione in rete (e che in genere offrono possibilità di ricerca, indirizzi e–mail, abbonamenti gratuiti, previsioni del tempo, notizie, ecc.).

6 La posta elettronica

6.1 Cos'è e come funziona

Quello della posta elettronica, o *e-mail* (*electronic mail*) è una dei servizi più diffusi e praticati su Internet. Molti servizi possono essere visti in analogia con quelli offerti dal servizio di posta tradizionale (talora indicata, su Internet, come *snail mail*, posta "lumaca"), anche se talvolta non ci sono delle corrispondenze precise.

Per cominciare, una *casella postale* è un luogo ove alcuni messaggi di posta elettronica vengono accodati. In molti casi, si tratta semplicemente di file di testo a cui vengono via via aggiunte le righe corrispondenti ai nuovi messaggi arrivati. In sistemi più sofisticati si fa uso di apposite basi di dati.

Molto spesso esiste una relazione univoca tra una casella postale ed un *indirizzo di posta elettronica*, che è l'indicazione in forma standard di come inviare un messaggio alla casella (ad esempio, l'espressione *mario@aziendaxyz.it* indica la casella di posta elettronica dell'utente *Mario* presso il sistema di posta elettronica della rete *aziendaxyz.it*).

Vi sono però delle eccezioni. Ad esempio:

- se in nell'azienda Mario si occupa delle vendite, si potrebbe pensare di creare un *alias* in modo tale che tutta la posta destinata a *vendite@aziendaxyz.it* venga direttamente sistemata nella casella di posta elettronica di Mario; ad una sola casella corrispondono due indirizzi.
- a volte tutta la posta elettronica di un determinato dominio (*unacosa@aziendaxyz.it*, *unaltracosa@aziendaxyz.it*, ecc.) confluisce in un'unica casella di posta elettronica (cosiddetta *casella multidrop*)
- ad alcuni indirizzi di posta elettronica non corrisponde una casella di posta elettronica ma un programma, che esegue dei compiti in funzione del messaggio ricevuto

Il simbolo @ (in inglese *at*, in italiano spesso letto "chiocchiolina") serve a separare il nome dell'utente dal nome del calcolatore che ospita la casella di posta dell'utente. Per le reti che ospitano più calcolatori appartenenti ad un dominio di secondo livello, normalmente è sufficiente indicare quest'ultimo, poiché il servizio di risoluzione dei nomi del sistema DNS è in grado di indicare quale calcolatore effettivamente si occupa della posta elettronica all'interno del dominio. In pratica, nonostante la casella postale di Mario sia fisicamente disponibile presso il calcolatore *posta.aziendaxyz.it*, non sarà necessario indicare come indirizzo *mario@posta.aziendaxyz.it* e sarà sufficiente *mario@aziendaxyz.it*.

6.2 Programmi per la posta elettronica

Dal punto di vista dei programmi coinvolti, quello della posta elettronica è uno dei servizi più complessi di Internet. Una buona parte della complessità deriva dal fatto che il servizio è stato concepito per poter funzionare anche con calcolatori che non sono permanentemente collegati ad Internet.

6.2.1 La spedizione

Teoricamente, se io dovessi spedire un messaggio alla casella *mario@aziendaxyz.it*, potrei chiedere al mio programma di posta elettronica di contattare il calcolatore che si occupa della posta presso la rete *aziendaxyz.it* e consegnare il messaggio.

Ma le cose non sono così semplici. Cosa succederebbe infatti se quel calcolatore al momento fosse fuori servizio (perché spento, sovraccaricato, in manutenzione o altro)? La spedizione non potrebbe avvenire. Per questo motivo è stato inventato un meccanismo di tipo *store and forward* (letteralmente, "immagazzina e inoltra"), che consiste nell'avere dei server per la posta in uscita a cui il nostro calcolatore si rivolge quando

6 La posta elettronica

deve spedire i messaggi. Il server accetta i nostri messaggi e se ne prende cura fino a quando la spedizione non avviene con esito positivo o risulta impossibile anche dopo diversi tentativi. La cosa interessante è che nel frattempo il mio calcolatore non deve preoccuparsi di nulla (un po' come, dopo aver spedito una lettera all'ufficio postale tradizionale, posso tranquillamente tornarmene a casa).

Nel malaugurato caso la spedizione non potesse andare a buon fine, il server di posta ci potrebbe avvisare con un messaggio (ad esempio, *user unknown*, "utente sconosciuto").

La spedizione avviene tramite un protocollo particolare, chiamato *SMTP*, ovvero *Simple Mail Transfer Protocol*, "protocollo semplice per il trasferimento della posta". Il programma che si occupa della spedizione dei messaggi si chiama *server SMTP*.

6.2.2 La lettura tramite prelievo dal server

Quand'è che il messaggio si considera consegnato? Non quando arriva al calcolatore dell'utente al quale è indirizzato, nè tanto meno quando viene effettivamente letto. Si considera consegnato quando viene accodato alla casella di posta elettronica a cui si fa riferimento con l'indirizzo.

Qui entra in gioco la ricezione dei messaggi. Infatti, non è affatto detto che il nostro utente Mario utilizzi per il proprio lavoro o per la gestione della posta proprio il calcolatore dove è presente la sua casella; a volte è così, ma è un caso piuttosto raro. Tipicamente, Mario dovrà invece indicare al suo programma di posta di farsi mandare tutti i messaggi che giacciono nella casella di posta dal calcolatore che la ospita.

Per recuperare la posta si utilizza normalmente il protocollo *POP3* (*Post Office Protocol, version 3*). Il programma che si occupa di consegnare la posta al legittimo proprietario con il protocollo POP3 si chiama *server POP3*.

Una tipica "conversazione" tra programma di posta e server POP3 potrebbe assomigliare alla seguente:

```
P: utente mario
S: ok
P: password segreto
S: ok, hai 2 messaggi per un totale di 4658 byte
P: mostrami la lista
S: messaggio 1 di 1240
S: messaggio 2 di 3418
P: mandami il numero 1
S: ecco (...)
P: mandami il numero 2
S: ecco (...)
P: cancella il numero 1
S: cancellato
P: cancella il numero 2
S: cancellato
P: ho finito
```

6.2.3 La lettura dei messaggi lasciati sul server

A volte i messaggi non vengono prelevati dal server ma lasciati sul server di posta. Ciò è comodo soprattutto quando la propria postazione di lavoro è variabile (a volte uso un calcolatore, a volte un altro), o quando si vuole dare un'occhiata alla posta da un calcolatore diverso dal solito.

Per queste esigenze, si possono usare un protocollo specifico, *IMAP* (*Internet Message Access Protocol*, protocollo per l'accesso ai messaggi Internet), oppure degli appositi servizi web (ossia server web configurati per consentire di leggere e scrivere messaggi di posta elettronica tramite un *browser*, i cosiddetti sistemi di *webmail*).

6.3 Anatomia di un messaggio email

Un messaggio di posta elettronica si compone di un'*intestazione* e di un *corpo*. L'intestazione contiene informazioni strutturate (una serie di coppie nome–valore in un formato specifico) su chi ha inviato il messaggio, a chi è destinato, l'oggetto dello stesso, ecc. Il corpo del messaggio contiene il testo vero e proprio del messaggio.

Dal momento della spedizione all'arrivo alla casella di destinazione, il messaggio viaggia accompagnato da una "busta" (*envelope*) con ulteriori informazioni.

6.3.1 Gli allegati

I messaggi di posta elettronica sono sempre di puro testo. È tuttavia possibile allegare file binari (immagini, documenti formattati con un elaboratore di testo, fogli elettronici, ecc.) grazie a opportuni meccanismi di conversione (in opera "dietro le quinte").

Normalmente viene utilizzato un sistema chiamato *MIME* (*Multipurpose Internet Mail Extension*, estensione con molteplici scopi per la posta su internet), che prevede diversi tipi di conversione a seconda delle necessità (tra questi, di frequente è usata una conversione chiamata *Base64* in cui tre byte del file originario vengono convertiti in quattro caratteri di un testo semplice). Storicamente, si sono utilizzati anche altre modalità, di cui talora si trova traccia, quali *uuencode/uudecode* (specialmente in ambiente Unix) e *BinHex* (specialmente in ambiente Macintosh).

Un particolare tipo di allegato è quello che rende il messaggio simile ad una pagina web; viene effettuata una conversione di ciò che scriviamo in HTML. Sebbene questa modalità sia spesso impostata come *default* in molti programmi per la gestione della posta, è bene non abusarne, poiché:

- i messaggi diventano inutilmente più grandi
- non tutti hanno programmi in grado di gestirli correttamente
- spesso messaggi siffatti portano problemi in fatto di sicurezza
- alcuni sistemi di gestione automatica della posta potrebbero non funzionare correttamente

Alcuni esempi di messaggi di posta elettronica con dei commenti sul modo in cui sono gestiti gli allegati sono disponibili nella mia [raccolta di esempi](http://www.tissino.it/esempi/email/) [<http://www.tissino.it/esempi/email/>].

7 I programmi per la posta elettronica

7.1 Il concetto di Mail User Agent

Per gestire la posta elettronica si usano programmi appositi (salvo si usino sistemi di *webmail*) chiamati **MUA**, ossia *Mail User Agent*. Esistono molti programmi in questa categoria, che si differenziano per interfaccia e ricchezza di funzionalità, nonché per licenza d'uso.

Molti programmi permettono l'accesso a caselle remote di posta elettronica tramite protocollo POP3 o IMAP, mentre alcuni (ad esempio *Pine* e *Mutt*), soprattutto in ambiente Unix/Linux, permettono l'accesso esclusivamente alla casella di posta locale del sistema, per cui si deve fare affidamento su altri programmi (come *Fetchmail*) per scaricare i messaggi.

7.2 Funzioni tipiche di un programma di posta

Un programma per la gestione della posta permette:

- Composizione *off-line* di messaggi
- Gestione di un indirizzario (una rubrica)
- Connessione al server SMTP per la spedizione
- Accesso alla casella di posta elettronica (locale o tramite POP3 o IMAP)
- Evidenziazione dei messaggi nuovi / non letti
- Ordinamento dei messaggi in base a vari possibili criteri
- Funzioni "Rispondi", "Rispondi a tutti", "Inoltra", "Reindirizza", "Spedisci di nuovo"
- Gestione di più cartelle per la posta arrivata e spedita (es. per argomento o per interlocutore)
- Gestione degli allegati
- Stampa dei messaggi ricevuti / spediti

Inoltre è bene che un programma di posta permetta:

- Filtraggio dei messaggi da scaricare (in base a dimensione, mittente, oggetto, ecc.)
- Separazione automatica dei messaggi ricevuti (con sistemazione nelle diverse cartelle)
- Gestione di più caselle di posta
- Autocomposizione di messaggi in uscita

7.3 La composizione di un messaggio

Quando si vuole comporre un messaggio, un programma di posta, al di là dell'interfaccia utente, presenta in genere uno spazio in cui inserire i destinatari del messaggio, l'oggetto, il testo del messaggio e l'indicazione degli allegati.

Se si tratta di un programma che gestisce più *account*, presenta anche la possibilità di scegliere con quale degli account si vuole operare.

7.3.1 I destinatari

Un messaggio di posta elettronica può avere uno o più destinatari. I destinatari possono essere indicati in tre diversi tipi di campi:

- nel campo **To:** (A:) si dovranno inserire i destinatari principali del messaggio
- nel campo **Cc:** (Cc:) si dovranno inserire i destinatari di copie per conoscenza del messaggio
- nel campo **Bcc:** (Ccn:) si dovranno inserire i destinatari di copie per conoscenza "nascoste"

7 I programmi per la posta elettronica

È bene sottolineare l'importanza della possibilità di inviare messaggi per conoscenza "nascosta". Chi riceve un messaggio può vedere nell'intestazione del messaggio gli indirizzi dei destinatari che, come lui, hanno ricevuto il messaggio. Se si vuole inviare un messaggio a più destinatari (ad esempio, a tutti i clienti o a tutti i propri conoscenti) senza che questi vengano a conoscenza vicendevolmente degli indirizzi, si dovrà far uso del campo *Bcc* (che spesso non appare immediatamente tra quelli disponibili).

7.3.2 La netiquette

Le regole di buon comportamento in rete (applicabili specialmente, ma non solo, a posta elettronica e newsgroups) vanno sotto il nome di *netiquette*. Eccone qualcuna piuttosto importante:

- la comunicazione scritta può essere causa di equivoci; è per questo che, ad esempio, è bene indicare esplicitamente, attraverso gli *smileys*, che si scrive qualcosa per scherzo.
- scrivere in MAIUSCOLO equivale a gridare
- evitare di spedire messaggi non sollecitati (cosiddetto *spamming*) a persone di cui si conosce l'indirizzo
- non pretendere / aspettarsi risposte immediate
- dare conferma della ricezione di un messaggio
- evitare allegati ingombranti

7.4 Funzionalità avanzate di un programma di posta elettronica

7.4.1 Aspetti organizzativi

Quando i messaggi e-mail scambiati cominciano ad essere molti, vale la pena di organizzarsi un po'. È bene prendere confidenza con:

- cartelle
- rubrica di indirizzi e gruppi
- filtri e regole

7.4.2 Account

Quando si configura un programma di posta elettronica di solito si impostano i seguenti parametri:

- indirizzo e-mail
- nome per esteso
- server per la posta in uscita (SMTP)
- server per la posta in arrivo (POP3 o IMAP), nome utente e password
- indirizzo a cui si chiede vengano inviate le risposte

Con il termine *account* si intende normalmente un indirizzo di posta elettronica.

Esistono però delle situazioni in cui più account possono essere definiti per uno stesso indirizzo di posta elettronica (ad esempio, in uno scenario in cui varia il nome per esteso o l'indirizzo per le risposte).

Inoltre, una stessa persona può disporre di diversi indirizzi di posta elettronica (ad esempio, per separare la posta privata da quella del lavoro, oppure perché abbonato a diversi provider). In questi casi è normale definire più account nel proprio programma di gestione della posta.

Se si usano più account di posta per i diversi provider, si deve tenere presente che il server SMTP di ciascuno è diverso. Molti provider non accettano di trasmettere messaggi che non sono inviati da clienti di altre reti (*we do not relay*, non ritrasmettiamo). Ciò comporta che non posso connettermi ad Internet con *provider1.it* e

spedire messaggi rivolgendomi a *smtp.provider2.it*.

7.5 Aspetti relativi alla sicurezza

Ci sono alcune considerazioni da fare in merito alla posta elettronica per un uso sicuro della stessa. Di seguito propongo alcuni spunti di riflessione.

Innanzitutto, si consideri che i messaggi di posta elettronica viaggiano "in chiaro" e che i contenuti quindi sono leggibili da tutti coloro che hanno la possibilità di accedervi (ad esempio, gli amministratori dei server attraverso i quali essi transitano); si eviti quindi di trasmettere dati sensibili e si prenda in considerazione l'ipotesi di usare tecniche crittografiche per la gestione della corrispondenza.

In secondo luogo, si tenga sempre presente che tutte le intestazioni dei messaggi possono essere falsificati in maniera molto semplice. Se ricevete un messaggio che apparentemente è stato spedito anche a qualcun altro (perché vedete un altro indirizzo oltre al vostro nel campo *To*: o nel campo *Cc*:) non è assolutamente detto che il messaggio sia stato effettivamente spedito all'indirizzo indicato.

Viene richiesto di conoscere una password per l'accesso alla posta elettronica disponibile in una casella di posta elettronica, ma non viene richiesta nessuna password per la spedizione dei messaggi, e quindi è relativamente facile spedire messaggi a nome di altri. Se ricevete un messaggio che apparentemente arriva da qualcuno, non potete avere la certezza che il mittente sia proprio quello indicato (a meno che non si faccia uso di firme digitali).

Tramite la posta elettronica, è semplice inoltrare i messaggi ricevuti. Esistono molte cosiddette "Catene di Sant'Antonio" (anche a fin di bene, a volte, ma in genere con effetti poco piacevoli). In particolare, bisognerebbe stare attenti alle "bufale" (*hoaxes*) che si diffondono su Internet in maniera molto veloce. Si dia un'occhiata ad esempio all'ottimo servizio antibufale [<http://www.attivissimo.net/#antibufala>] di Paolo Attivissimo (consiglio vivamente di leggere gli arretrati).

7.6 Servizi utilizzabili tramite posta elettronica

Tramite la posta elettronica si può anche:

- richiedere e ricevere pagine web e file
- farsi ricordare impegni ed appuntamenti
- farsi avvertire della variazione di una pagina web
- abbonarsi a *mailing lists*
- inviare messaggi a *newsgroups*
- (inviare e) ricevere fax
- inviare SMS

8 Domande ed esercizi

Queste domande e questi esercizi sono pensati come spunti di discussione e di riflessione su quanto appreso. Al momento sto riflettendo sull'opportunità di scrivere anche le risposte. [Nei corsi che tengo, preferisco darle a voce dopo un piccolo "dibattito".]

8.1 Domande

1. In generale, è la stessa cosa indicare come indirizzo `http://www.qualchesito.it` e `http://qualchesito.it`? Giustifica la risposta e porta, se possibile, degli esempi a sostegno di essa.
2. Se il server DNS risolve il nome `www.esempio.org` di un calcolatore con l'indirizzo IP `68.13.45.90`, è la stessa cosa indicare l'URL `http://www.esempio.org` e `http://68.13.45.90`? Giustifica la risposta e porta, se possibile, degli esempi a sostegno di essa.
3. Una nota casa automobilistica, che non citerò, fa comparire informazioni pubblicitarie corredate da un indirizzo internet nella forma `buy@casadellauto.com`. Di che tipo di indirizzo si tratta?
4. In una pagina web in cui è richiesto che inserisca in due campi di un modulo il mio nome utente e la mia password per accedere ad un determinato servizio (ad esempio, di webmail), posso evitare di compilare il modulo inserendo nome utente e password direttamente nell'URL, secondo lo schema `http://utente:password@www.host.tld`? Sì? No? Perché?

8.2 Esercizi ed attività pratiche

1. Leggere la guida per le ricerche con Google [<http://www.google.it>] e con Altavista [<http://www.altavista.it>]
2. Il browser che usi permette di compiere tutte le operazioni elencate nel capitolo sui browser?
3. Il programma di posta elettronica che usi permette di compiere tutte le operazioni elencate nel capitolo sui browser?
4. Per la gestione del tuo account di posta elettronica hai la possibilità di accedere ad un servizio di webmail? Se sì, ci sono dei vantaggi e degli svantaggi nell'accesso via web rispetto all'accesso tradizionale con un programma di posta elettronica per la gestione della posta fuori linea?

Qualcuno ha suggerimenti per altre domande di questo tipo?

9 Licenza d'uso per questo documento

9.1 GNU Free Documentation License

Version 1.2, November 2002

Copyright (C) 2000,2001,2002 Free Software Foundation, Inc.
59 Temple Place, Suite 330, Boston, MA 02111-1307 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

9 Licenza d'uso per questo documento

The "Cover Texts" are certain short passages of text that are listed, as Front–Cover Texts or Back–Cover Texts, in the notice that says that the Document is released under this License. A Front–Cover Text may be at most 5 words, and a Back–Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine–readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard–conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine–generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front–Cover Texts on the front cover, and Back–Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the

9 Licenza d'uso per questo documento

publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- **A.** Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- **B.** List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- **C.** State on the Title page the name of the publisher of the Modified Version, as the publisher.
- **D.** Preserve all the copyright notices of the Document.
- **E.** Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- **F.** Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- **G.** Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- **H.** Include an unaltered copy of this License.
- **I.** Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- **J.** Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- **K.** For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section,

9 Licenza d'uso per questo documento

and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

- **L.** Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- **M.** Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- **N.** Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- **O.** Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License,

9 Licenza d'uso per questo documento

provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.