

ProGeAs

Telematica per le arti e lo spettacolo

I n t e r n e t

Internet

- Storia e evoluzione

DARPA SECONDO DOPOGUERRA

GUERRA FREDDA

STATI UNITI

ISTITUZIONI

MILITARI

CONTROLLO

Internet

- Attraverso il proprio computer domestico, per mezzo di un semplice modem collegato al telefono, è possibile trasmettere e ricevere ogni genere di dati, informazioni sotto forma di testi, immagini, suoni. Si possono visitare in modo virtuale i maggiori musei del mondo e si possono consultare, esplorare biblioteche e banche dati.
- In base a ciò il sapere universale è e sarà disponibile a ciascuna persona, in qualsiasi momento, in qualsiasi parte del mondo.

Internet

- Il combinarsi del computer con le reti telematiche (collegamento a computer remoti) permette il potenziamento delle prestazioni multimediali da esso offerte. L'uso delle reti fa accedere ad un sistema di informazioni ed idee dalle dimensioni e dalle caratteristiche prima sconosciute e fa interagire con altri individui in maniera veloce ed economica (e-mail, news group).
- All'utente di Internet si spalancano le porte di un mondo di notevole ricchezza: attraverso reti locali (LAN Local Area Network all'interno della scuola) o in aree più vaste (WAN Wide Area Network co-partecipazione a progetti condivisi di più scuole).

Internet – la storia

- Nel 1957 il dipartimento della Difesa degli Stati Uniti d'America costituì **l'Advanced Research Projects Agency (ARPA)**, un'agenzia di ricerca istituita per ristabilire la superiorità degli USA nel campo scientifico e tecnologico in risposta al successo del lancio sovietico del primo satellite orbitale Sputnik.
- Gli studi dell'ARPA erano concentrati sulla conquista dello spazio, i missili balistici intercontinentali, gli esperimenti nucleari. Solo incidentalmente vennero investite risorse nel campo delle comunicazioni telematiche. La **costituzione di una rete di comunicazione (ARPANET - 1969)** era infatti funzionale a una miglior gestione delle risorse, visto che dava la possibilità di collegarsi con i centri di ricerca esterni e con i loro elaboratori elettronici.

Internet – la svolta

- La vera rivoluzione in questo ambito fu **l'annullamento del paradigma per cui per comunicare in rete fosse necessario un sistema centrale di smistamento.**

Internet – una struttura diffusa

- Un'idea di base era una ragnatela di collegamenti in cui i "pacchetti" di dati, partendo da un computer d'origine, girassero indipendentemente per vie decise di volta in volta da ogni singolo computer (nodo) su cui transitavano, per poi incontrarsi alla fine del viaggio e riassemblarsi nella forma originaria.
- Un sistema di trasporto "intelligente", capace di risolvere problemi come la perdita dei dati e, in caso di traffico eccessivo o caduta dei collegamenti, capace di trovare strade alternative per far giungere a destinazione le informazioni.

Internet – distribuzione a pacchetto

- Un sistema che aveva nella sua stessa natura "diffusa" la garanzia di sicurezza:
 - nessun nodo era infatti fondamentale di per sé, poteva saltare o essere posto sotto controllo, in ogni caso il rischio era che solo una piccola parte delle informazioni venisse intercettata, una parte insignificante per chi la rilevava e la cui mancanza non avrebbe compromesso la decodifica dell'intero messaggio, giunto per altre vie a destinazione.

Internet è nata

- E così l'ARPA raccolse tutte queste idee innovativi, le tecnologie relative e nel **1969 lanciò la nuova rete chiamata ARPANET**, collegando quattro nodi dislocati nel sud-ovest degli Stati Uniti (UCLA, Stanford Research Institute, University of California Santa Barbara, University of Utah).

Internet può essere regolata?

- Per rendere l'idea di libertà e apertura insiti nella rete Internet, va sottolineato che, nonostante esistano varie authority per la corretta gestione e sviluppo della rete, **non c'è un nodo centrale finanziario o gestionale.**
- Le organizzazioni esistenti (per la maggior parte sconosciute agli utenti, autofinanziate e tutte paritariamente coordinate) sorreggono, plasmano, ma non dirigono, né controllano lo sviluppo di questo insieme multiforme di reti, ognuna con la propria storia e peculiarità.
- Per questa ragione, oltre che per l'estensione globale della rete, **Internet travalica ogni ordinamento giuridico territorialmente limitato** ed è animata da una sorta di spirito anarchico, libertario e auto-regolante, che la differenzia da tutti gli altri media.

Che cos'è Internet?

- 1. **un network di reti**, ovvero un insieme di *protocolli condivisi, router e circuiti telematici*;
- 2. **una comunità di persone** che usa e sviluppa *quelle stesse reti per incontrarsi e comunicare*;
- 3. **un insieme di risorse accessibili attraverso quelle reti**, ovvero *pagine Web, posta elettronica, download di file (FTP), forum di discussione, servizi, informazioni, commercio elettronico*.

Internet – triplice natura

- Internet è un insieme di reti che collega tra loro milioni di computer in tutto il mondo, per questo è chiamata anche la "Rete delle Reti".
- Per rete si intende un qualunque collegamento che unisca tra loro due o più computer; collegamento che può realizzarsi tramite la linea telefonica, le nuove “fibre ottiche”, o i recentissimi collegamenti wireless o satellitari.

Come funziona Internet

- Per scambiare le informazioni i computer hanno bisogno di “parlare” nella medesima lingua, cioè di capirsi l’un l’altro.
 - Questo è reso possibile dall’utilizzo di quello che viene definito un “protocollo di comunicazione” cioè un insieme di convenzioni, di regole comuni, che ogni macchina è in grado di accettare, comprendere e utilizzare: questo protocollo è noto col nome abbreviato di TCP/IP ovvero *Transfer Control Protocol/Internet Protocol*.
- Tale protocollo ha la possibilità di dividere le informazioni in tanti piccoli pacchetti che verranno poi spediti alla medesima destinazione. Una volta raggiuntala, i pacchetti verranno ricomposti in modo da ricreare l’informazione completa.

Internet non è il WWW !

- Il World Wide Web è oggi il servizio più famoso offerto dalla Rete, tanto da essere diventato sinonimo di Internet stessa. Eppure la Ragnatela Mondiale – è questa la traduzione di World Wide Web – è nata tardi, agli inizi degli anni Novanta, dalle ricerche portate avanti da alcuni studiosi del CERN di Ginevra.
- La caratteristica più famosa e appariscente del WWW, che ne ha determinato l'immediato, indiscusso e planetario successo, è la possibilità di creare e utilizzare gli ipertesti.
- Un ipertesto è un nuovo modo di organizzare le informazioni dando la possibilità all'utente di "saltare" da un documento ad un altro, e da un testo ad un filmato o ad un immagine, attraverso dei collegamenti che in gergo tecnico sono detti "*link*", costruendo quindi un percorso non legato ad alcuno schema predefinito.

Che cos'è un browser

- E' ormai molto in voga il termine “navigare in Internet” ed in effetti entrare nella Rete ha molte caratteristiche in comune con la navigazione: ci si trova in un mare sconosciuto nel quale all'inizio è difficile muoversi e trovare ciò che ci interessa.
- Ma ci sono degli strumenti che ci vengono validamente in aiuto: un browser è appunto questo, un programma che permette di orientarci in questo enorme arcipelago conducendoci felicemente in porto.
- Il browser (che in italiano potremmo tradurre con "navigatore") è lo strumento che ci permette di visualizzare le pagine Web dei siti che ci sono su Internet.
- I due browser più famosi, entrambi distribuiti gratuitamente sono:
 - Netscape Navigator
 - Microsoft Internet Explorer.

Come si usa un browser

- Per pervenire al porto (sito) che desideriamo non dobbiamo far altro che digitare nella casella *Location* (posizionata sulla barra degli strumenti) l'URL (indirizzo Internet) del sito di destinazione (che comincia di solito per: *www.ecc...*) e cliccare su *OK*. Oppure possiamo cliccare sul menù *File* quindi *Apri* e digitare l'indirizzo del sito che ci interessa.
- Gli altri tasti del browser assomigliano a quelli del comune registratore che abbiamo a casa:
 - *Back o Indietro*: serve per tornare alla pagina precedentemente visualizzata
 - *Forward o Avanti*: porta alla pagina seguente
 - *Home*: porta alla pagina di partenza
 - *Print o Stampa*: manda in stampa la pagina al momento visualizzata

I motori di ricerca

- Un'informazione "disordinata" del web pone di fronte a due tipi di problemi: la **reperibilità e l'affidabilità**.
- Per quanto riguarda il primo problema, una soluzione può essere quella di affidarsi agli strumenti di ricerca disponibili in Rete che sono i **motori di ricerca per termini e gli indici sistematici**.
- Per navigare in Internet, cioè per raggiungere il sito che desideriamo, dobbiamo conoscere l'indirizzo, digitarlo nell'apposita casella chiamata "*Location*" e premere il tasto *Invio*.
- Qualora però non lo conoscessimo abbiamo la possibilità di utilizzare degli strumenti appositi che "corrono" ininterrottamente attraverso il World Wide Web catalogando tutte le pagine che incontrano. Questi programmi sono i Motori di Ricerca.
- In un'apposita casella è possibile digitare uno o più termini chiave e chiedere al programma di effettuare una ricerca, terminata la quale compariranno sullo schermo gli indirizzi di tutti i siti che contengono quella parola e che noi potremo comodamente visitare cliccando su di essi.

Due tipi di motori di ricerca

- Esistono due tipi di motori di ricerca:
- *Per termini*: per effettuare una ricerca è sufficiente digitare una o più parole chiave nell'apposita casella La **ricerca attraverso un motore di ricerca per termini** è molto comoda nel caso di **nomi** propri o nel caso che le informazioni che cerchiamo siano caratterizzabili con termini molto specifici.
- Occorre tuttavia tener presente che si tratta di una ricerca meccanica: il programma cioè si limiterà a cercare i termini che noi gli abbiamo fornito all'interno di un immenso indice alfabetico in suo possesso. Se l'input dato al motore è l'argomento che ci interessa, l'output sarà una lista di documenti presentati in ordine di rilevanza.
- A questa tipologia di motori appartengono
 - **Google** (www.google.it) e
 - **Altavista** (www.altavista.it).

- *Per “argomenti”*: la ricerca si svolge scegliendo tra le categorie disponibili quella che si avvicina di più al nostro caso. La **ricerca sistematica** avviene su cataloghi ragionati di risorse, suddivisi e organizzati, le cosiddette **web directories**.
- La ricerca quindi non è più meccanica bensì è il risultato di una decisione umana. Un catalogo ragionato di questo genere si basa su una sorta di "albero del sapere" che bisogna percorrere partendo da categorizzazioni più generali per arrivare via via verso quelle più specifiche.
- Tra i Motori di Ricerca del secondo tipo uno dei più completi è *Yahoo* il cui indirizzo è www.yahoo.it
 - Yahoo! e Lycos non sono gli unici servizi di catalogazione sistematica delle risorse disponibili in Internet, ma attualmente sono quelli più completi e sono disponibili anche in versione italiana.

Affidabilità dell'informazione in rete

- Per quanto riguarda **l'affidabilità** dell'informazione un consiglio generale è questo: una volta trovata la pagina informativa che reputiamo interessante è bene risalire alla home page del sito che la ospita (utilizzando le apposite icone attive disponibili su molte pagine). In questo modo avremo informazioni su chi ha immesso in Rete quell'informazione, in quale contesto ecc.

Gli strumenti

- Lo scambio di file, lo scambio di informazioni, lo scambio di idee, lo scambio di parole. **Quella allo scambio è la vera vocazione di Internet.**
- Oggi, di fatto, esistono infiniti modi per comunicare con altri utenti, garantendo sempre ottimi risultati
 - instant messaging (invio di messaggi in tempo reale)
 - Email
 - Scambio di file (documenti, musica, video, ecc.)
 - chiamate vocali (voice over IP)
 - teleconferenze
- Questo **aumento nel volume degli scambi** si sposa con la **moltiplicazione del pubblico e dei punti di vista.**

Il problema della sicurezza

- Internet è un sistema complesso e molto sofisticato che comprende tecnologie e interventi umani. A fronte delle considerazioni fatte sulla vastità del web e l'eterogeneità del materiale che viene pubblicato, è bene tener presente alcune norme di sicurezza
- Da una parte è opportuno **equipaggiare il proprio sistema operativo in modo da evitare accessi non autorizzati.**
 - Microsoft Internet Explorer e Netscape Navigator, per esempio, sono dotati di caratteristiche potenti che garantiscono livelli di sicurezza differenziati, ossia una soglia di protezione più alta per i siti non noti e una più bassa per i siti fidati.
 - In alternativa è possibile installare un software di protezione del sistema operativo e di controllo dell'accesso a Internet, una specie di controllore che vigila su quelli che sono i movimenti "sotterranei" durante la navigazione.
- Attenzione: non si tratta di un intervento umano, il controllo è tecnico e/o preventivo, ma non discrimina i contenuti dei siti su cui si naviga.

Internet come strumento di ricerca

- Come fonte di informazione, Internet da una parte apre le porte a nuovi contenuti, dall'altra rende facilmente disponibili, anche se in una forma nuova, le fonti tradizionali:
 - Biblioteche,
 - Banche dati,
 - Musei virtuali,
 - Gallerie,
 - Giornali,
 - Libri elettronici, ecc.
- Tutti sappiamo come funziona una biblioteca eppure utilizzarla al meglio, servirsene utilizzandone tutte le potenzialità non è facile, né banale. Lo stesso discorso vale per Internet dove non basta fornire una parola-chiave ai motori di ricerca.
- Internet, lo abbiamo già sottolineato, mette a disposizione informazione di natura molto eterogenea, raggiungibile attraverso canali diversi.

- oltre a sapere ***che cosa vogliamo trovare,***
dobbiamo sapere che tipo di informazione
vogliamo trovare.

Due tipi di informazione

- differenza fondamentale ***tra informazione strutturata e informazione "libera"***.

Informazione strutturata

- Attraverso Internet sono dunque accessibili, fra molte altre cose, banche dati specializzate, contenenti informazioni fortemente strutturate come il catalogo di una biblioteca, atlanti stellari, informazioni geografiche ecc. Le consulteremo quando ci servono informazioni di quel tipo.

Informazione libera

- non tutte le informazioni che ci servono sono rintracciabili presso banche dati strutturate come il catalogo di una biblioteca e **quindi dovremo ricorrere all'informazione "libera" disponibile in Rete.**
- Cosa significa? Significa che istituzioni come musei o gallerie d'arte, critici d'arte, appassionati di storia e storici, viaggiatori appassionati e enti del turismo, **hanno ritenuto utile rendere accessibili informazioni attraverso pagine web.**
- **In breve, anziché un'informazione strutturata, troveremo un'informazione eterogenea e anche la nostra sarà una ricerca diversa.**

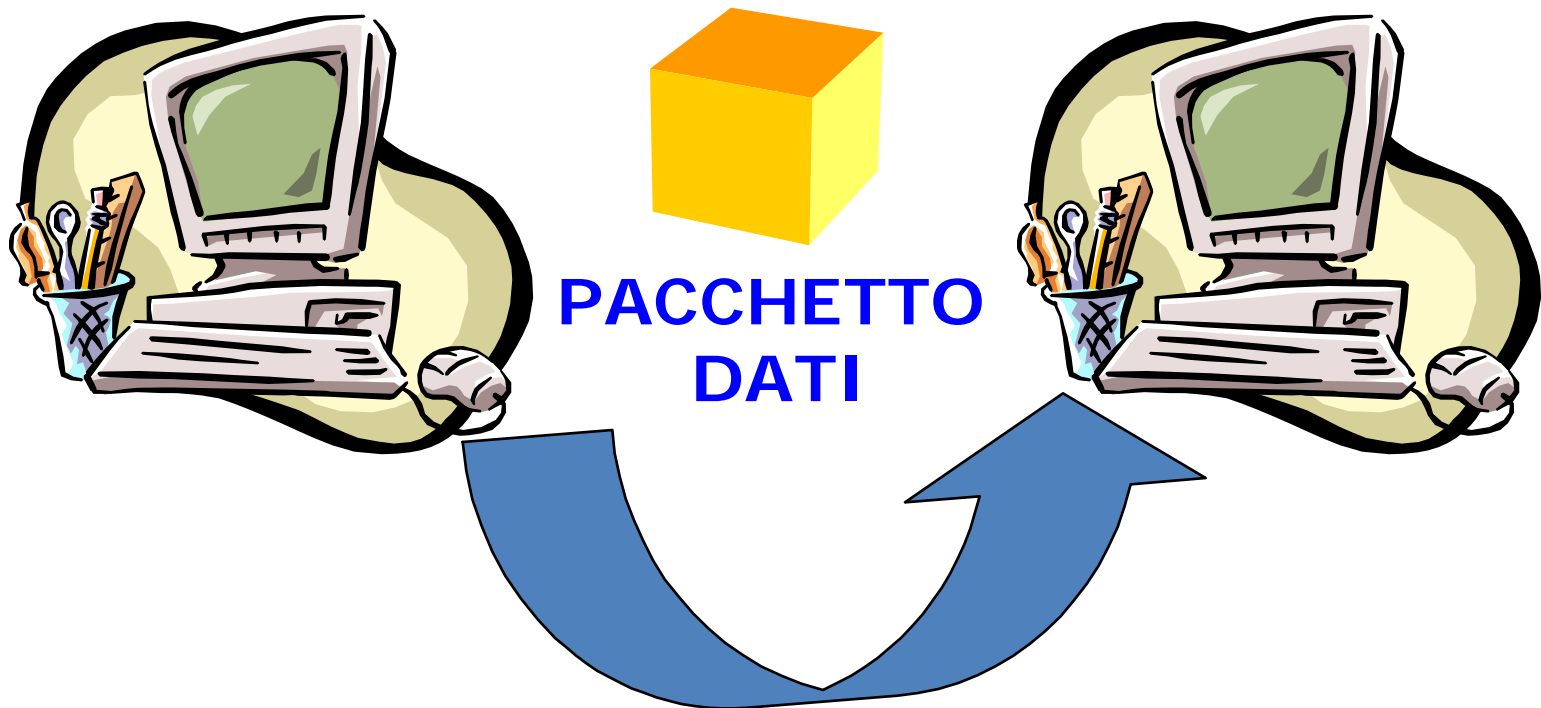
Informazione ordinata o disordinata?

- Abbiamo visto la differenza fra **informazione "ordinata"** e **informazione "disordinata"**. Entrambe presentano vantaggi e difficoltà:
- **informazione "ordinata"** di una banca dati è più puntuale e affidabile ma, pur essendo raggiungibile attraverso Internet, non è integrata nella grande ragnatela del Word Wide Web.
- **'informazione "disordinata"** di Word Wide Web è più difficile da valutare dal punto di vista dell'affidabilità e spesso per accedervi siamo costretti a navigazioni lunghe e qualche volta frustranti. **E' però un'informazione sterminata.**

Internet

- I PROTOCOLLI

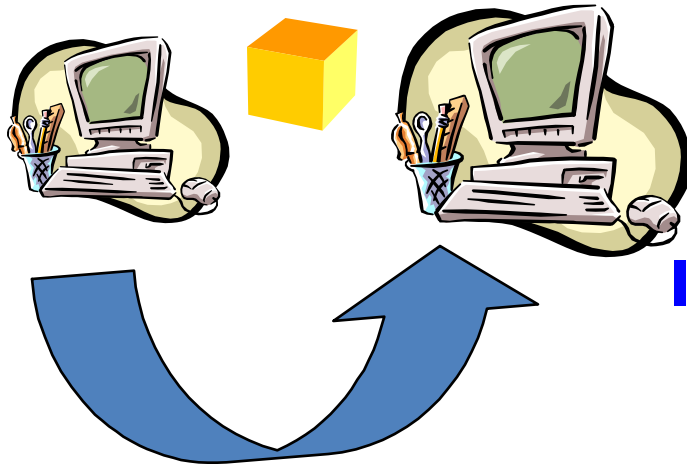
PROTOCOLLO=REGOLE COMPORTAMENTALI



Internet

- I PROTOCOLLI

PROTOCOLLO=REGOLE COMPORTAMENTALI



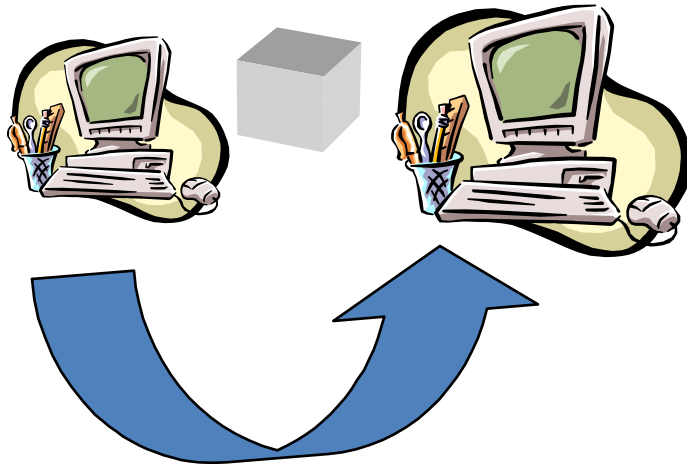
**PACCHETTO
DATI
VIENE SCOMPOSTO
IN TANTE PARTI E INVIATO**

**IL PROTOCOLLO RICOMPONE L'ORDINE DEI
DATI e ne
RICOSTRUISCE LA SEQUENZA**

Internet

- I PROTOCOLLI

TCP = Trasmission Command Protocol



IP = Internet Protocol

WWW

- Un documento di testo che contiene dei collegamenti (“***link***”) ad altri documenti è detto un ***ipertesto***.
- Un semplice documento di testo risiedente su un *host* è scandito (letto) sequenzialmente dall’inizio alla fine.
- Un ipertesto viene letto sequenzialmente fino a quando si giunge a un link, a questo punto si può decidere di attivare il link e la lettura prosegue su un nuovo ipertesto, eventualmente risiedente su un altro *host*.

Navigazione

- Di qui si può poi riprendere l'ipertesto precedente dal punto di interruzione, saltare ad altri documenti mediante altri link e così via. Questo modo di procedere si chiama ***navigazione***.
- Per istituire un sistema di ipertesti occorre che:
 1. gli ipertesti siano codificati in un opportuno formato, che preveda i link;
 2. che gli ipertesti siano gestiti da appositi server che soddisfino alle richieste dei clienti, ossia degli utenti che vogliono accedere all'informazione.

- Il formato usato da Berners-Lee per codificare gli ipertesti si chiama ***Hyper Text Markup Language (HTML)***, il protocollo di accesso che corre fra cliente e server si chiama ***Hyper Text Transfer Protocol (HTTP)***.
- Questa architettura si è presto imposta ed è venuta a costituire l'applicazione distribuita di maggior successo in Internet.
- L'insieme degli *host* che contengono la base dati costituisce la cosiddetta “ragnatela che ricopre il mondo”: il World-Wide Web appunto, o semplicemente ***Web***.

Indirizzamento nel Web: URI e URL

- Gli ipertesti non costituiscono l'unica forma di documentazione esistente in Internet. Basi dati di altro tipo, soprattutto quelle più vecchie preesistenti all'introduzione degli ipertesti, usano altri formati.
- Queste basi dati risiedono su *host*; su ognuno di questi *host* corre un processo chiamato *information server* per accedere al quale sono spesso stati definiti specifici protocolli di livello applicativo.

- Il Web ha lo scopo di comprendere il maggior numero possibile di informazioni. Possiamo quindi pensare il Web come costituito da **oggetti** ai quali si accede con un numero di protocolli indeterminato e ampliabile (si deve dare la possibilità di introdurre nuovi tipi di informazione, gestiti da nuovi tipi di *information server* a cui si accede con protocolli ancora da inventare).
- Per navigare il Web occorre un programma detto “**Web browser**”. Esso è uno *User Agent* che interagisce con i diversi *information server* utilizzando gli appropriati protocolli.

URI e URL

- Ogni protocollo ha di solito le proprie convenzioni per identificare gli oggetti su cui opera.
- Il più generico identificativo (ossia il nome) che si usa per indicare un certo oggetto, indipendentemente dal tipo dell'oggetto e dal protocollo d'accesso, viene designato come ***Universal Resource Identifier (URI)*** ed è definito nella RFC 1630.
- Uno ***Uniform Resource Locator (URL)***, definito in RFC 1738) è una particolare forma di URI corrispondente a uno degli attuali protocolli d'accesso. Di qui si vede che in pratica nei casi comuni gli identificatori che interessano sono sempre degli URL.

- Un URL è un “*locator*” e ha quindi la funzione di “*locate*”, ossia individuare una **risorsa** fornendone un’identificazione astratta.
- Esso è costituito da una stringa di caratteri strutturata come segue:

<schema>:<scheme-specific-part>

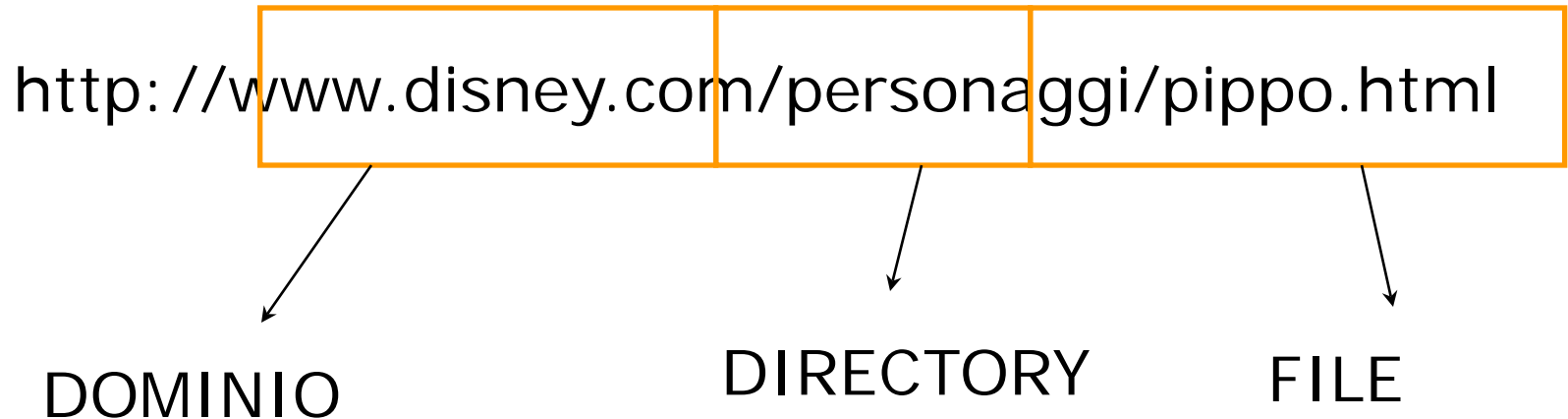
dove <schema> indica il metodo d’accesso (di solito un protocollo), ad esso segue il carattere “:” (due punti) e quindi una stringa che descrive la risorsa e il cui formato varia a seconda del metodo d’accesso.

<schema> è una stringa di caratteri alfabetici minuscoli.

Internet

- URL

Uniform Resources Locator



URL con schema file

- Uno schema un po' speciale è lo schema **file**. Con questo schema l'URL assume la forma:
file://<host>/<path>
dove <host> è il *domain name* di un computer sul quale il percorso <path> risulta accessibile.
- Di solito l'accessibilità non è possibile se non utilizzando un certo protocollo, ma ad esempio una risorsa sull'*host* locale su cui corre il *browser* è senz'altro accessibile senza protocolli, per cui il seguente URL dà accesso al file locale definito dal percorso C:\temp\stazioni.txt1:
file://localhost/C:/temp/stazioni.txt
- Il nome localhost è sottinteso, può quindi mancare e l'URL precedente diventa:
file:///C:/temp/stazioni.txt

URL con schema http

- A parte lo schema ***file***, tutti gli altri schemi indicano un protocollo della famiglia TCP/IP.
- Se la risorsa voluta è un ipertesto il protocollo per accedervi è ***http***. Un URL con schema http ha la struttura:
http://<host>:<port>/<path>
- *<host>* è il *domain name* oppure l'indirizzo IP dell'*host* su cui corre il server http;
- *<port>* è il numero di porta tcp per accedere al server http. Se la stringa
- *“:<port>”* è omessa si usa il valore di *default* che è 80;
- *<path>* è un *http selector*, ossia un'indicazione data al server http per localizzare la risorsa all'interno dell'*host*.

- Se la stringa “<path>” è omessa si può omettere anche il carattere “/” finale.
- In questo caso la risorsa è una risorsa convenzionale definita all’atto dell’installazione del server http, spesso un file html di nome index.html (la *HomePage* del server) e che risiede in un directory specifico dell’*host*, noto al server e non visibile dal cliente.
- Ad esempio i seguenti URL:
 - http://195.75.86.68:80
 - http://195.75.86.68
 - http://195.75.86.68/
 - http://www.it.ibm.com
 - http://www.it.ibm.com/index.htmlsono tutti equivalenti e danno accesso alla stessa risorsa (la *HomePage* di IBM Italia).

URL con schema ftp

- Un URL con schema ftp serve ad accedere a risorse (file e directory) su un host mediante il protocollo ftp e ha la struttura:
- `ftp://<user>:<password>@<host>:<port>/<path>`
- Questo percorso è su un computer DOS/Windows e quindi usa il carattere “\” per separare i campi.
- Negli URL il percorso usa sempre il carattere di separazione “/”, conformemente alla convenzione di UNIX.

- • <user> è il *login name* sull'*host*. Se omesso questo campo equivale al valore "anonymous";
- • <password> è la relativa *password*. Se si usa accesso di tipo anonymous, come password si deve indicare l'indirizzo di *e-mail* dell'utente.
 - Se è omessa e il server la richiede, il *browser* la deve chiedere all'utente;
- • <host> è il *domain name* oppure l'indirizzo IP dell'*host*;
- • <port> è il numero di porta tcp per accedere al server. Se la stringa ":<port>" è omessa il valore di *default* è 21;
- • <path> è il percorso ftp.
 - Se omesso corrisponde al *directory* di base del server ftp.
 - Diversamente equivale a dare una serie di comandi *cd* (*change directory*) a partire da tale *directory*.

- I tre URL seguenti sono equivalenti:

ftp://152.2.254.81

ftp://sunsite.unc.edu

ftp://anonymous:alazzari%40magoo.cedi.unipr.it@sunsite.unc.edu:21/

- Tutti danno accesso al *directory* di base dell'*anonymous server* sull'*host* "sunsite.unc.edu".
- Si noti nella terza variante la codifica del campo <password>.
- In questo campo secondo le regole si sarebbe dovuto mettere "alazzari@magoo.cedi.unipr.it", tuttavia il carattere "@" è un carattere speciale nell'URL: esso serve a separare <password> da <host>.
- Scrivendo in tale modo, il *browser* avrebbe interpretato erroneamente:
<password>=alazzari <host>=magoo.cedi.unipr.it@sunsite.unc.edu
- Per far capire al *browser* che il primo carattere "@" incontrato nell'URL non è un separatore ma è parte della stringa <password> lo sostituisce dalla sua codifica ASCII espressa in esadecimale e preceduta dal carattere "%".