

Fondamenti di Informatica

Accademia di Belle Arti di Verona

Università degli Studi di Verona

A.A. 2016-2017

Docente - Vincenzo Giannotti

CAPITOLO 5 – RETI DI COMPUTER e INTERNET

Reti di computer

I sistemi di elaborazione dell'informazione si sono evoluti, negli ultimi decenni, verso una integrazione tra le risorse di elaborazione stesse e i sistemi di telecomunicazione (ICT).

Utilizzare il computer, per la maggioranza delle persone, significa «**comunicare**» con altre persone, aziende, istituzioni: scambiare documenti, fare acquisti on-line, chattare, richiedere certificati etc..

Questo settore dell'informatica si chiama «**telematica**», termine che rende conto della integrazione tra le due tecnologie: quella dell'informatica e quella delle telecomunicazioni.

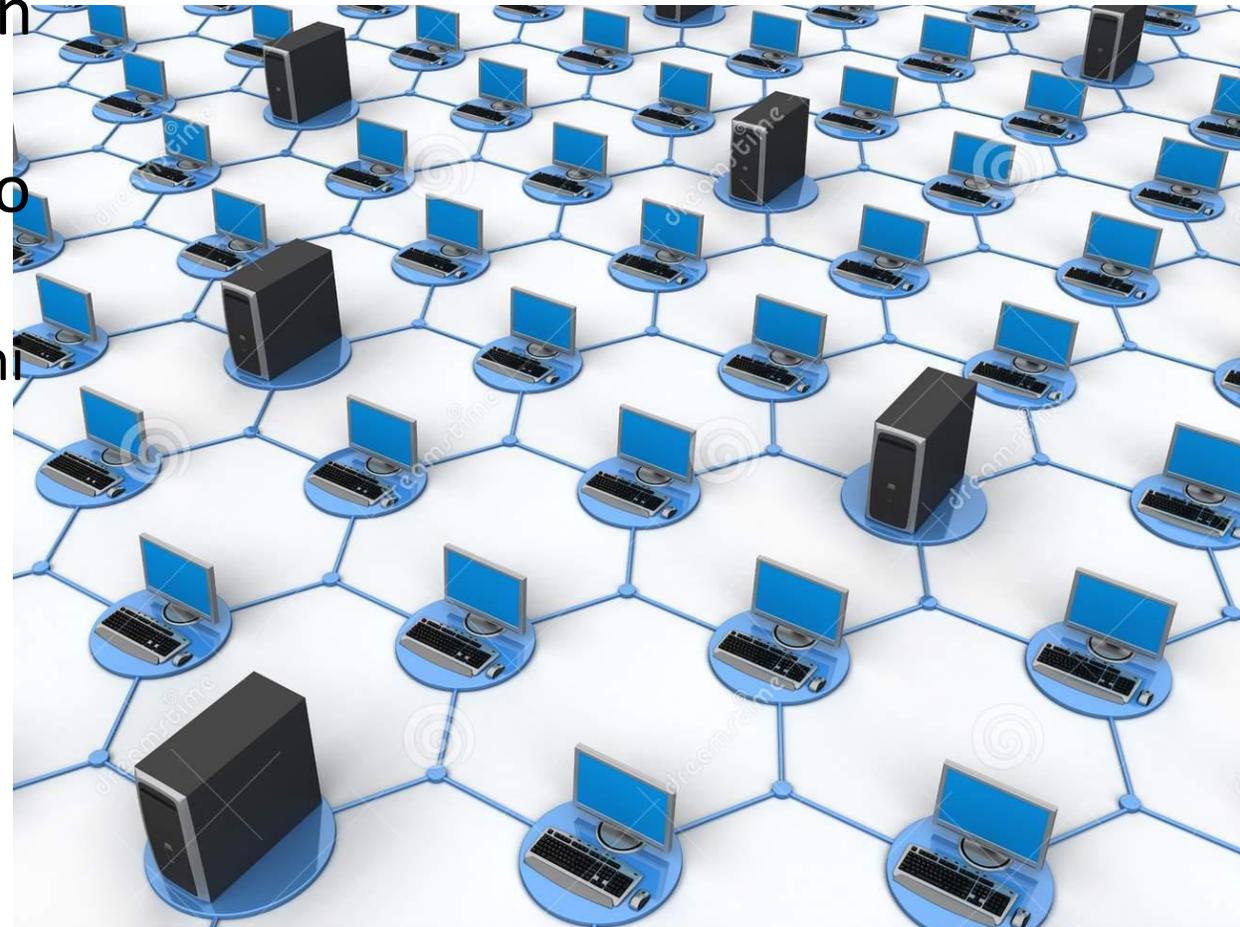
L'aspetto principale e centrale della telematica riguarda le

Reti di computer

Reti di Computer

Una rete (Network) di computer è un insieme di sistemi di elaborazione messi in grado di comunicare tra loro per consentirci di:

- Condividere software e applicazioni
- Consultare e gestire archivi
- Scambiare dati
- Elaborare informazioni distribuite su più macchine
- Gestire grandi progetti
-



Reti di Computer

L'impiego delle tecnologie telematiche consente di poter comunicare senza doversi preoccupare della distanza tra i diversi sistemi di elaborazione: l'estensione della rete dipende solamente dalle esigenze che un dato soggetto ha di entrare in contatto con altri soggetti.

Per questo esistono reti che operano a livello locale (**LAN** – Local Area Network), per esempio all'interno di una azienda, ovvero reti che operano ad un livello geografico più ampio (**MAN** – Metropolitan Area Network e **WAN** – Wide Area Network), per esempio per connettere una città o diverse sedi di una stessa azienda.

Inoltre una rete può essere aperta (**open**) o chiusa (**closed, proprietary**).

Come funziona una Rete di Computer

Protocolli di Comunicazione

Nella comunicazione tra persone dobbiamo seguire delle regole se vogliamo farci capire. Per prima cosa quando ci esprimiamo, sia verbalmente, sia in forma scritta, dobbiamo fare in modo che il nostro discorso rispetti le regole fondamentali della «**sintassi**»; questo significa che quando vogliamo esprimere un concetto dovremo farlo in maniera intelligibile, con frasi costruite secondo delle regole che il nostro interlocutore possa facilmente decodificare.

VADO A MANGIARE NONNA

VADO A MANGIARE, NONNA

Una virgola può salvare una vita

Usa la punteggiatura, salva la nonna

Protocolli di Comunicazione

Un'altra cosa di cui tenere conto è l'aspetto «**semantico**» del discorso: se parliamo di «reti» con un pescatore, presumibilmente stiamo discorrendo di pesca piuttosto che di informatica. All'interno di un discorso su un dato argomento, non possiamo inserire degli incisi che parlano d'altro. Alcuni lo fanno, ma non va bene.

Uguualmente, poiché anche la comunicazione non verbale può entrare a far parte del discorso, se stiamo esprimendo un concetto che trasmette gioia, sarebbe bene che la nostra espressione non esprimesse disgusto.

Infine quando parliamo con altri è buona norma «**sincronizzare**» il nostro discorso col loro: se parliamo tutti nello stesso momento la comunicazione è inutile. Normalmente, anche solo per buona educazione, interveniamo solo quando gli altri hanno terminato di parlare.

Protocolli di Comunicazione

Nella comunicazione informatica valgono più o meno le stesse regole e queste regole si chiamano «**protocolli**».

I protocolli dunque servono a definire la sintassi, la semantica e la sincronizzazione della comunicazione tra computer e possono essere realizzati mediante sistemi hardware, sistemi software, o sistemi misti hardware e software.

In linea generale i protocolli definiscono:

- **Il formato di dati.** Nella comunicazione informatica vengono scambiate stringhe di bit. Se queste eccedono una data lunghezza prestabilita sono suddivise in «pacchetti» di dimensioni adeguate. Normalmente il messaggio è diviso in due parti: l'intestazione e i dati. L'area di intestazione contiene i campi più rilevanti per il protocollo.
- **Gli indirizzi del mittente e del destinatario.** Gli indirizzi sono utilizzati per identificare il mittente e il destinatario previsto. Gli indirizzi vengono memorizzati nell'area di intestazione.
- **Routing.** Quando mittente e destinatario non sono collegati direttamente ma attraverso sistemi di comunicazione intermedi, i messaggi debbono poter essere correttamente inoltrati. In Internet per esempio le reti sono collegate tramite dispositivi intermedi detti Router. Questa modalità in cui più reti sono poste in comunicazione tra loro, si chiama **internetworking**.

Protocolli di Comunicazione

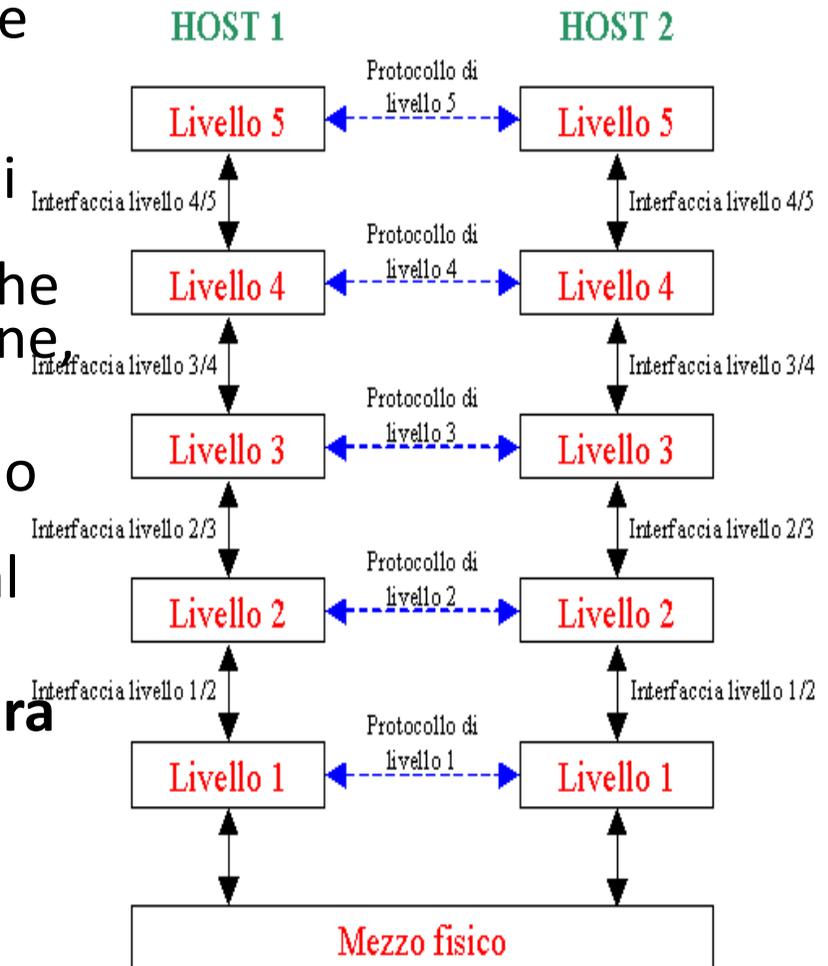
- **Rilevamento di errori di trasmissione.** Un requisito irrinunciabile nelle comunicazioni in rete è che ne sia garantito un funzionamento privo di errori. Per ottenere questo risultato alla fine dei pacchetti vengono aggiunti dei cosiddetti CRC (Cyclic Redundancy Check), rendendo possibile per il ricevitore rilevare differenze causate da errori.
- **Perdita di informazioni - timeout e tentativi.** I pacchetti potrebbero non giungere a destinazione o arrivare con eccessivo ritardo. In alcuni protocolli, un mittente può richiedere che sia confermata la corretta ricezione entro un certo periodo di tempo. Allo scadere del tempo il pacchetto viene considerato perduto e viene ritrasmesso.
- **Direzione del flusso di informazioni.** La direzione del flusso deve essere dichiarata nel caso vi siano trasmissioni che, per esempio, possono verificarsi solo in una direzione alla volta. Il protocollo deve possedere delle regole che governano le modalità di accoglimento di una richiesta di comunicazione, nel caso in cui le due parti vogliano ottenere il controllo allo stesso tempo.
- **Controllo della sequenza.** Abbiamo visto che quando i messaggi sono troppo lunghi, vengono suddivisi in pacchetti che sono spediti indipendentemente. Per motivi diversi abbiamo anche visto che alcuni pacchetti possono perdersi in rete o ritardare o prendere strade diverse. Ciò può comportare che alcuni pacchetti arrivino “fuori sequenza”. Se nel frattempo un pacchetto è stato rispedito, si possono avere dei duplicati che ovviamente debbono essere eliminati.
- **Controllo del flusso.** Si tratta di una operazione necessaria quando il mittente trasmette più veloce rispetto al ricevitore o rispetto ad apparecchiature intermedie nella rete.

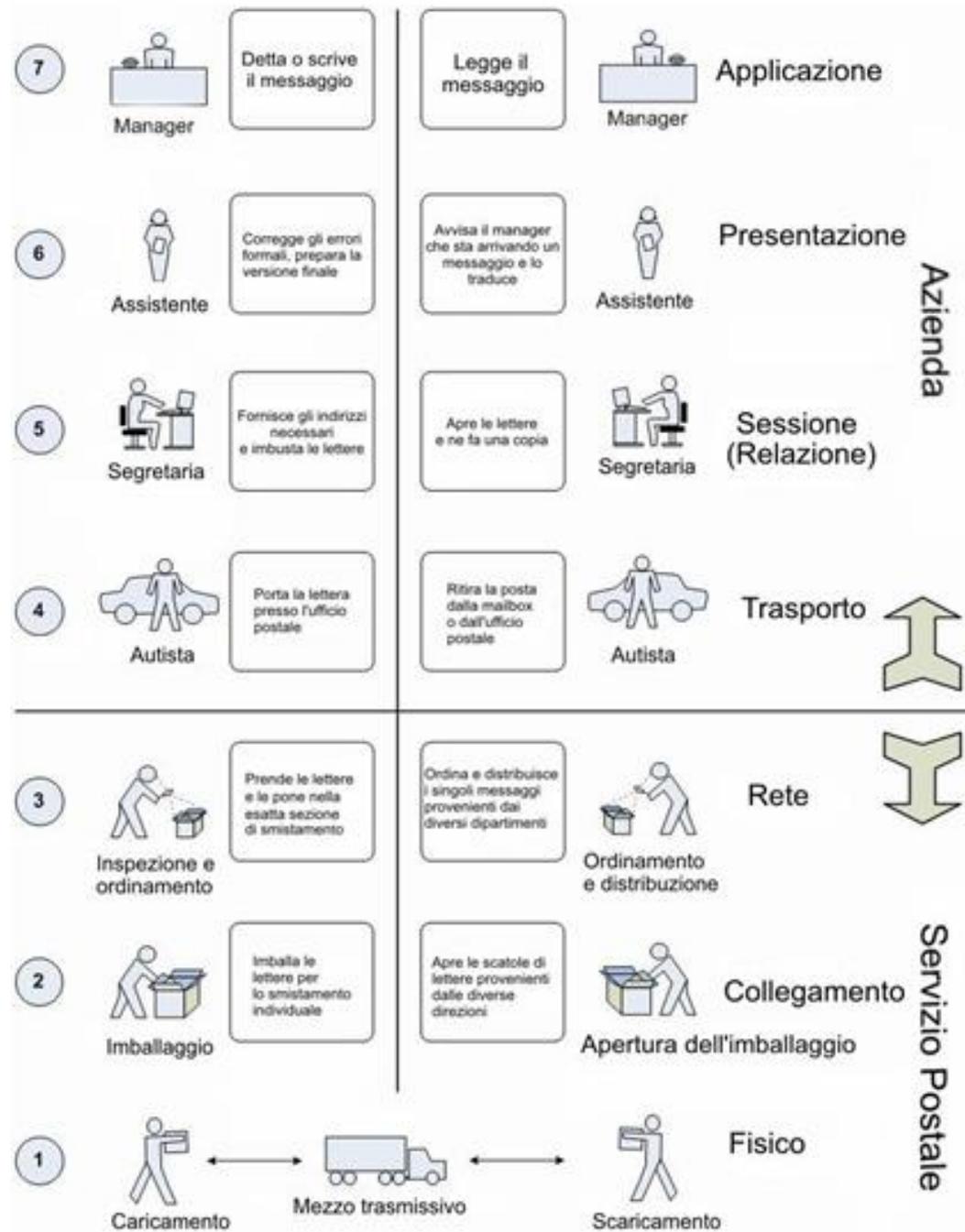
Architetture di Rete

Tutte le reti di telecomunicazione sono sempre organizzate per livelli. A partire dal livello più alto che è quello in cui il messaggio viene scritto, i dati passano ai livelli inferiori e vengono arricchiti ad ogni passaggio con le informazioni di controllo relative al protocollo del livello in cui si trovano. Una volta che il messaggio è arrivato al livello più basso, che è quello fisicamente connesso con la linea di comunicazione, esso viene effettivamente trasmesso.

In ricezione avviene il percorso inverso e ciascun protocollo analizza e riconosce le informazioni di controllo che gli competono, consentendo di ricostruire il messaggio fino al fruitore finale.

- Un insieme di livelli e di protocolli è chiamato **architettura di rete**.
- Il livello più basso è quello che risulta **fisicamente connesso** con la linea di comunicazione.



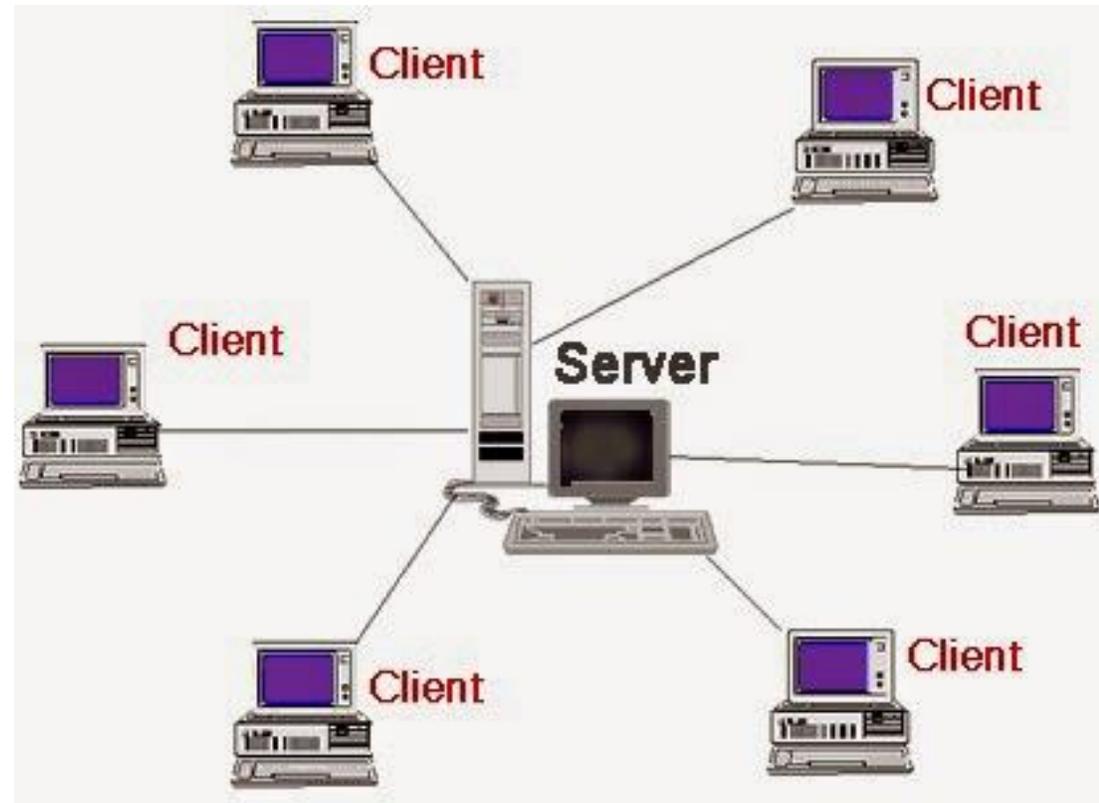


Parallelo tra una lettera e modello ISO - OSI

Il Modello Client-Server

A livello applicativo il modello più utilizzato per le reti è quello denominato **client/server**, dove la comunicazione avviene tra due differenti tipi di apparati:

- il **client** è il computer che inoltra la richiesta di servizi in rete
- il **server** è il computer (o il sistema di elaborazione dati) che funge da centro ospitante (**host** di archivi, software etc..) e che si occupa di distribuire i dati e le informazioni agli utenti della rete.



Il Modello Client-Server

Attraverso il Server più utenti possono **condividere risorse** di vario genere che possono essere variamente distribuite:

- Software applicativo
- Archivi di dati
- Stampanti e altre periferiche
- Computer particolarmente veloci
- Unità di archiviazione (dischi magnetici, dischi ottici, nastri ..).

Quando poi cominciamo ad interconnettere tra loro reti diverse, ecco che gettiamo le basi per «l'Internetworking» ossia **Internet**: la infrastruttura di rete che mette in collegamento tra loro computer, sistemi informatici e reti di computer di tutto il mondo.

The Internet

«**The Internet**», come la chiamano gli anglosassoni, ha avuto un impatto così sconvolgente nella società moderna che non è consigliabile procedere oltre senza prima approfondire alcuni aspetti del suo funzionamento generale e acquisire un minimo di terminologia. Internet infatti non è solo una infrastruttura di rete ma è anche un «luogo» e una «intelligenza collettiva» nella quale molti hanno trovato il loro *modus vivendi*.

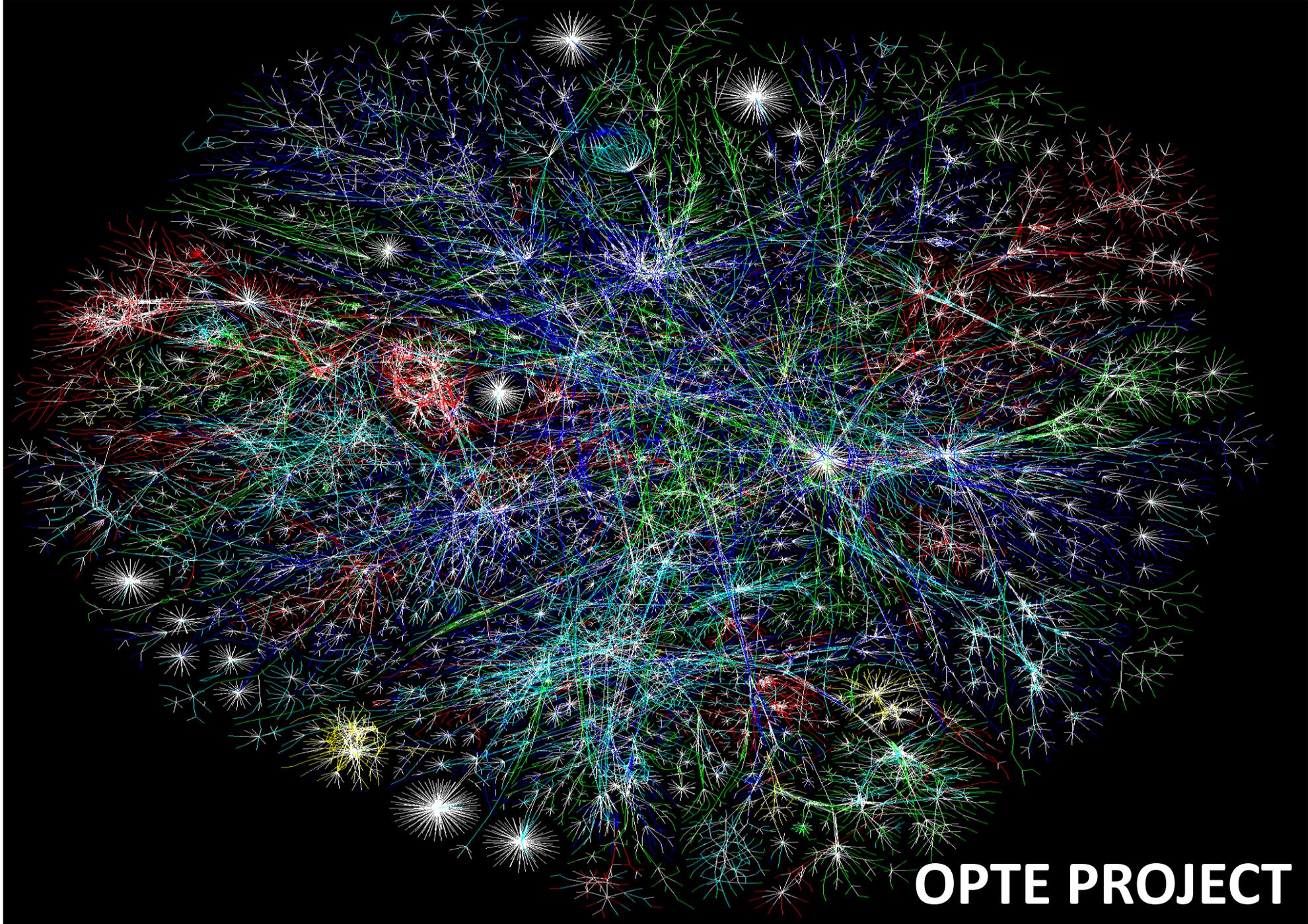
Come tutti noi siamo in grado di riconoscere le maggiori problematiche dell'ambiente, sappiamo cosa è una città e un condominio, sappiamo muoverci nel traffico urbano e quali sono le autorità che governano la nostra vita quotidiana, allo stesso modo dovremmo potere cavarcela con Internet e il Web.

Internet e il World Wide Web



Internet è un sistema aperto di reti di computer, distribuito sull'intero pianeta; una rete di reti alla quale chiunque può accedere (fatti salvi eventuali permessi) per colloquiare con altri computer. E non solo: è anche possibile colloquiare direttamente con altre persone connesse alla rete attraverso il loro computer o il loro smartphone.

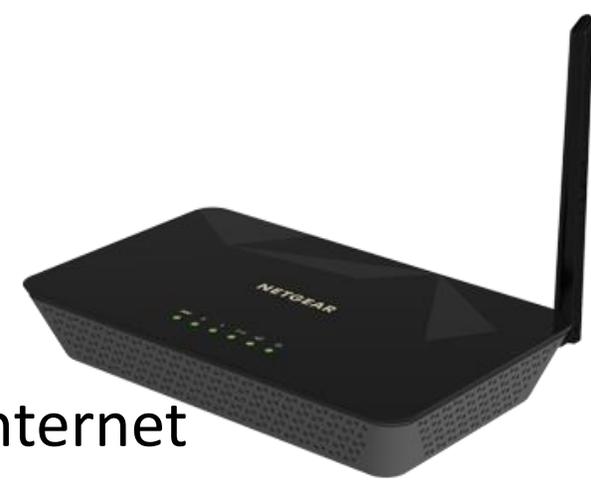
Il **World Wide Web** (web significa ragnatela) meglio conosciuto come **The Web** o **WWW**, è invece il principale servizio di Internet, che permette di navigare in un mare di informazioni utilizzando un sistema di **ipertesti** che collega tra loro risorse e servizi diversi distribuiti su tutta la rete (**link**).



OPTE PROJECT

Un po' di storia

Tre passi importanti



I più importanti passaggi tecnologici che aprono la strada a Internet sono tre:

1. L'invenzione del **MODEM** - uno strumento sviluppato nel 1949 per trasmettere segnali radar in ambito militare e che in seguito fu perfezionato per il progetto SAGE (Semi-Automatic Ground Environment), negli USA a partire dal 1958. Si tratta di un dispositivo elettronico ricetrasmittente che serve a far comunicare due dispositivi digitali (i.e. due computer) attraverso una linea di comunicazione analogica (il doppino del telefono). Ciò può avvenire realizzando due operazioni di conversione: la prima cosiddetta di «**modulazione**» del segnale digitale in analogico; la seconda di «**demodulazione**» del segnale analogico in digitale: da qui il nome mod-dem.

Tre passi importanti

2. Il secondo fondamentale passaggio fu dato dallo sviluppo del concetto di **Hyperlink**. Verso la fine degli anni '50 diversi studiosi cominciarono a teorizzare su come realizzare il concetto di *cross-reference* tra computer, ossia su come far sì che una data istanza di un documento potesse rinviare ad una informazione relazionata da qualche altra parte. Il termine hyperlink fu coniato intorno al 1965 da Ted Nelson, un pioniere dell'informatica, filosofo e sociologo, fondatore del progetto Project Xanadu, che aveva lo scopo di realizzare una rete di computer potendola gestire attraverso una semplice interfaccia. Questo stesso concetto fu ripreso dopo un paio di decenni dall'inventore del Web Tim Berners Lee, di cui parleremo più avanti.

Tre passi importanti

3. L'ultimo importante passaggio è legato alla introduzione del concetto di **Time-sharing**.
 - Time-sharing significa “condivisione del tempo” e nella pratica informatica significa che più utenti possono utilizzare nello stesso tempo una medesima risorsa. In termini più generali questo vuol dire che un gran numero di utenti possono interagire contemporaneamente con un singolo computer, anche se non ne possiedono uno personale.
 - Il concetto di time-sharing fu descritto per la prima volta da John Backus nell'estate del 1954 al MIT di Boston e in seguito da Bob Bemer in un suo articolo del 1958 dal titolo "How to consider a computer" nel quale scriveva: "The computers would handle a number of problems concurrently. Organizations would have input-output equipment installed on their own premises and would buy time on the computer much the same way that the average household buys power and water from utility companies."
 - Un **terminale** di computer è un dispositivo hardware elettronico (normalmente monitor e tastiera) che viene utilizzato per inserire dati in un computer centrale.

Come è nato Internet

Internet risale alla fine degli anni sessanta e fu il risultato di un progetto dell'ente governativo degli Stati Uniti ARPA (Advanced Research Project Agency). Il progetto in questione si chiamava **ARPANet**.

L'obiettivo del progetto, nato con finalità di difesa militare, era quello di realizzare una rete particolarmente solida che consentisse ai ricercatori di varie università USA di interconnettersi.

ARPANet si basava su una architettura innovativa che consentiva ai messaggi di transitare su percorsi diversi cosicché la distruzione di uno o più nodi (per esempio a causa di un attacco nucleare) non avrebbe compromesso la capacità di comunicazione del sistema.

Come è nato Internet

Nel 1973 essa fu aperta al NORSTAR (Norwegian Seismic Array) che fu il primo componente non USA ad entrare a far parte della rete. Dopo ciò, in pochi anni, ARPANet divenne molto popolare e nel 1977 la sua gestione «non militare» fu assunta dalla NSFNET (National Science Foundation NETwork); in quel periodo si potevano contare oltre 100 computer connessi alla rete.

Nel 1990 il lato militare di ARPANet fu definitivamente abbandonato e anche la restante rete passò in carico a NSFNET; in questo nuovo assetto la rete continuò ad espandersi soprattutto in Canada e in Europa.

Tuttavia, per consentire che ARPANet potesse espandersi a livello globale mancavano ancora un paio di eventi: tali eventi accaddero tra il 1991 e il 1992.

Come è nato Internet

Il primo evento fu l'apertura per fini commerciali della rete; nel 1990 infatti cominciava a farsi sentire la pressione di chi avrebbe voluto utilizzarla a fini commerciali, cosa a quel tempo proibita in quanto rete dedicata a scopi di ricerca. Di questa esigenza NSFNET prese atto l'anno seguente e nel marzo del 1991 mise mano alle proprie "Acceptable Use Policy" proprio per consentirne una utilizzazione anche a fini commerciali.

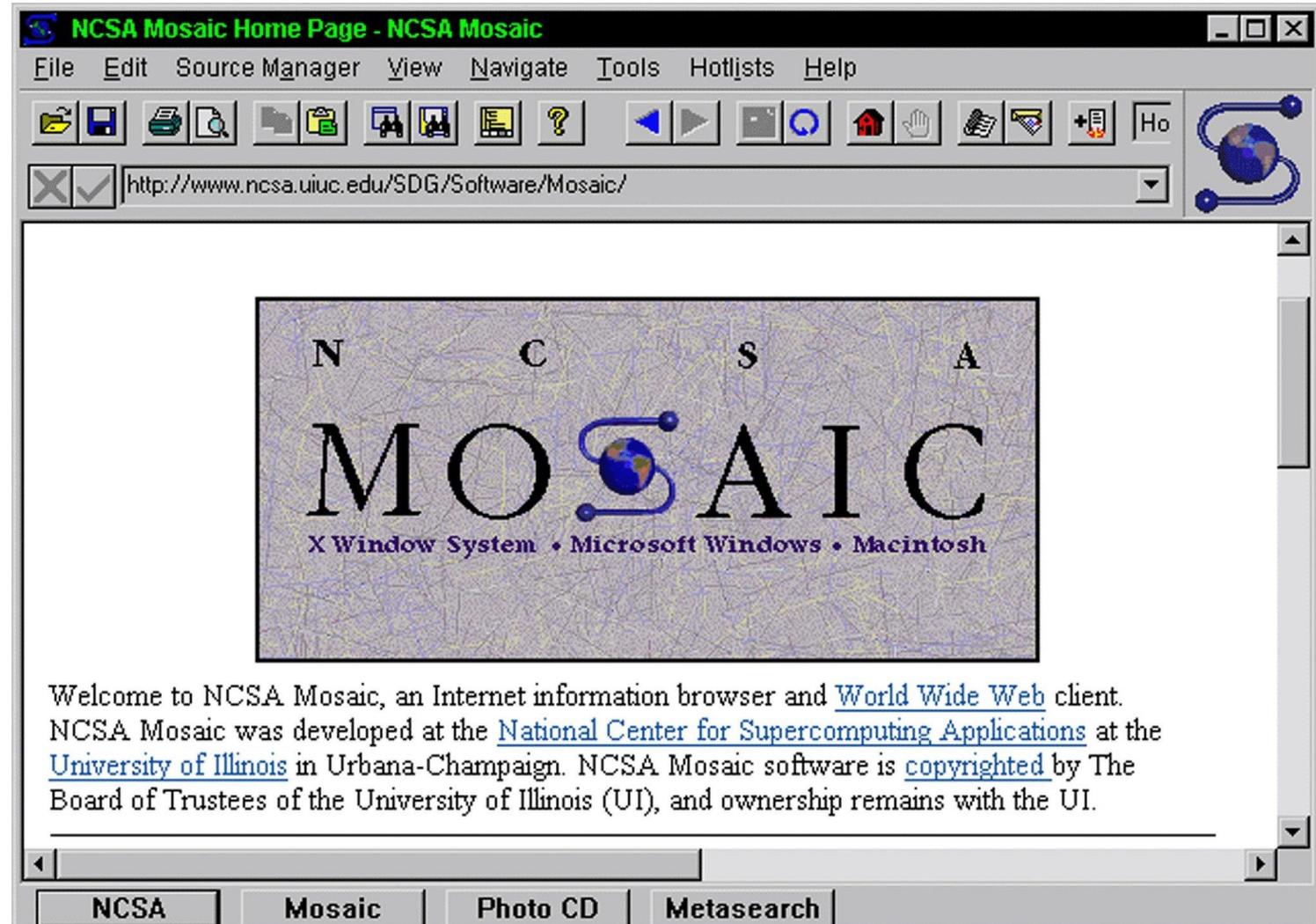
Il secondo evento fu l'invenzione, nello stesso anno, del World Wide Web, meglio noto con WWW o semplicemente Web, ad opera del ricercatore inglese Tim-Berners Lee che al CERN di Ginevra definì il protocollo HTTP (HyperText Transfer Protocol), un sistema che permette una lettura non-sequenziale dei documenti saltando da un punto all'altro mediante l'utilizzo di rimandi: gli hyperlink (o più semplicemente «link»). Il primo sito web al mondo fu realizzato da Lee quello stesso anno ed è consultabile a questo indirizzo:

<http://info.cern.ch/hypertext/WWW/TheProject.html>

Come è nato Internet

Due anni dopo, nel aprile del 1993, fu presentata la versione 1.0 di **Mosaic**, scritta dal National Center for Supercomputing Applications (**NCSA**).

Mosaic fu il primo browser moderno della storia, che di fatto aprì Internet anche al di fuori del mondo accademico e aziendale.



Come funziona Internet

Come funziona Internet

Quando veniamo assunti in una azienda, il primo giorno di lavoro ci vengono dati una sedia, una scrivania e un computer. Ci vengono presentati i colleghi, il nostro responsabile e probabilmente l'amministratore di sistema.

L'amministratore di sistema tra le altre cose ci fornisce una username e una password per accedere alla rete aziendale e a Internet. A questo punto, senza stare troppo a pensarci su, digitiamo i nostri codici sulla tastiera e da quel momento ci vengono aperte le porte di accesso a Internet.

Ma proviamo a chiederci: chi è che apre queste porte?

Gli Internet Service Provider

Le porte vengono aperte dagli **Internet Service Provider** o più semplicemente Provider o meglio ancora ISP.

Un ISP è un'organizzazione che fornisce servizi per l'accesso e l'uso di Internet; stiamo parlando di servizi come:

- l'**Accesso** che ci permette il collegamento alla rete utilizzando un computer o un dispositivo mobile o una rete informatica
- il **Transito** che consente ai nostri dati di attraversare una rete di computer e poi un'altra ancora e un'altra ancora, fino alla destinazione
- la **Registrazione di un Dominio** che ci consente di definire una nostra area autonoma di controllo all'interno della rete secondo le regole e le procedure del **Domain Name System** (DNS) di cui parleremo più oltre

Gli Internet Service Provider

- il **Web Hosting** quando abbiamo bisogno di pubblicare un nostro sito web accessibile attraverso il World Wide Web
- la **MailBox** (posta elettronica) per inviare e ricevere messaggi e documenti attraverso il computer, proprio come in un servizio postale.

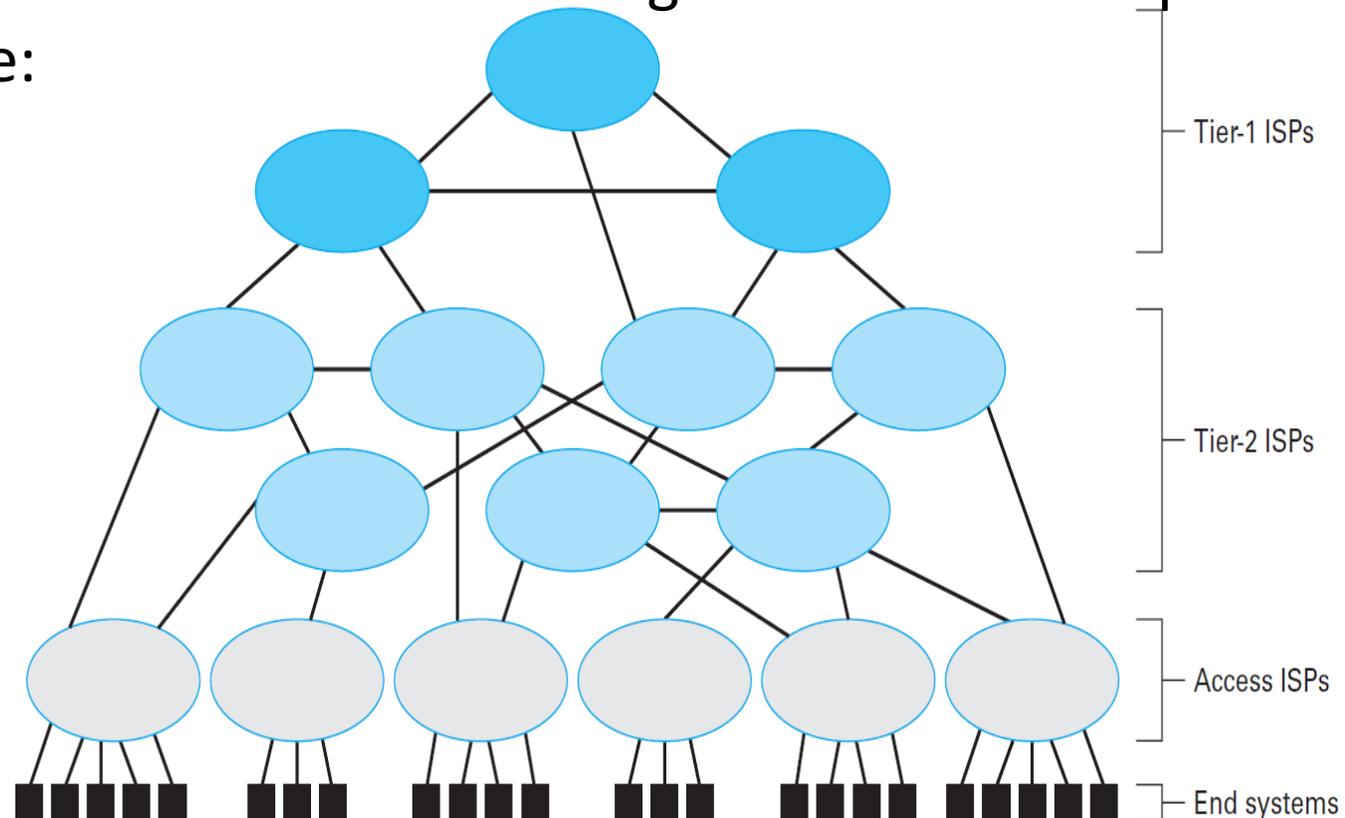
In base ai diversi servizi forniti possiamo avere diverse tipologie di ISP: Gli ISP più grandi garantiscono normalmente buona parte dei servizi di accesso, transito, MailBox etc.

altri ISP possono essere specializzati solo nella fornitura di alcuni di questi servizi: per esempio Yahoo! o Gmail forniscono servizi di MailBox ma non sono Access Provider, pur essendo dei colossi dell'informatica.

Gli Internet Service Provider

Nonostante Internet sia una rete di dimensioni planetarie, non vi è alcuna autorità che definisca il livello a cui operano gli ISP e tuttavia, pur in carenza di una ufficialità formale, è comunemente riconosciuta un'organizzazione del tipo rappresentato nella figura seguente:

Nella parte superiore di questa gerarchia sono pochi ISP di primo livello (**Tier-1**) che hanno accesso libero alla rete globale: essi non devono comprare il transito da altri provider e non pagano diritti di accesso.



Gli Internet Service Provider

Non è detto tuttavia che un Tier-1 possa accedere liberamente a tutta la rete e può essere che alcune destinazioni siano raggiungibili solamente con degli **accordi di peering** con altri provider. Tuttavia tali accordi, per i Tier-1, sono non onerosi in quanto il peering è una interconnessione volontaria tra reti che consente lo scambio di traffico tra gli utenti di ciascuna rete.

Gli ISP di secondo livello (i **Tier-2**) sono dei provider che necessariamente debbono appoggiarsi ad altre reti per poter raggiungere tutto Internet e ciò può avvenire con accordi di peering ovvero con accordi di transito, questi ultimi normalmente a pagamento.

Al livello **Tier-3** infine, si trovano quei provider che a loro volta acquistano il transito da altri provider di livello superiore.

Gli Internet Service Provider

Se consideriamo il fatto che questo tipo di organizzazione non è regolamentato da alcun ente, diventa comprensibile il fatto che non sempre si riesca ad attribuire il corretto livello ad un dato provider. Per fare una attribuzione corretta dovremmo infatti essere in grado di conoscere se quel Provider opera anche sulla base di eventuali accordi onerosi con altri Provider e questo non è sempre possibile.

Tra gli ISP che con buona probabilità possono essere classificati Tier-1 troviamo colossi come AT&T negli Stati Uniti, Deutsche Telekom AG, Telecom Italia Sparkle.

Tra quelli che operano invece a livello Tier-2 troviamo per esempio: British Telecom, Vodafone, Easynet e molti altri.

Come funziona Internet

Anche per Internet, come per tutte le reti, la cosa più importante da fare è fissare delle regole con cui condurre le attività di comunicazione. Tali regole come abbiamo visto, sono definite come **protocolli**. Un protocollo di rete serve a coordinare la trasmissione dei messaggi tra i diversi computer.

In Internet la trasmissione viene gestita col protocollo **TCP/IP** (Transport Control Protocol / Internet Protocol).

Questo protocollo si occupa di suddividere il messaggio in pacchetti che vengono trasmessi in rete (TCP) e instradati lungo il percorso migliore per consegnarli all'indirizzo del destinatario (IP). Una volta a destinazione i pacchetti vengono ricomposti nel messaggio originale (TCP).

Indirizzi IP e DNS

Abbiamo visto in precedenza che l'accesso a Internet può avvenire a diversi livelli. L'ultimo livello è quello in cui si trova l'utilizzatore, ossia un computer, uno smartphone o un altro dispositivo connesso alla rete e che ne utilizza i servizi.

Esattamente come per il servizio postale tradizionale, in cui il portalettere recapita la posta all'indirizzo specificato, anche in una rete di computer è necessario che ciascun utente abbia un proprio indirizzo che lo rende raggiungibile in maniera univoca. Quando parliamo della rete Internet questo indirizzo si chiama “**IP Address**” dal nome del protocollo che opera al relativo livello: **Internet Protocol**.

Indirizzi IP e DNS

L'Internet Protocol non è altro che un'etichetta numerica che viene assegnata a ciascun dispositivo che partecipa alla rete e che utilizza il protocollo Internet per le regole di comunicazione.

Un IP address ha due funzioni principali: una è quella di identificazione del dispositivo e un'altra è quella di indirizzamento. Questa duplice funzione è ben resa nel seguente passaggio delle specifiche del Internet Protocol "A **name** indicates what we seek. An **address** indicates where it is" Internet Protocol – DARPA Internet Program Protocol Specification (Settembre 1981).

Nella sua prima forma **IPv4** (disegnata nel 1983 per ARPANet) l'Internet Protocol fu definito come un numero a 32 bit in grado di indirizzare 2^{32} dispositivi.

Indirizzi IP e DNS

IPv4 può dunque individuare quattro miliardi di indirizzi. Quattro miliardi sono molti e infatti IPv4 è utilizzato ancora oggi tuttavia, considerando la velocissima crescita della rete, la disponibilità di indirizzi si sta velocemente esaurendo. Per questo motivo nel frattempo è stato realizzato il nuovo sistema di indirizzamento **IPv6** a 128 bit che dovrà rimpiazzare in via definitiva IPv4.

Gli IP Address sono rappresentati in una modalità di semplice lettura secondo la seguente notazione: **aaa.bbb.ccc.ddd** nel caso del IPv4.

L'indirizzo è dunque rappresentato da quattro gruppi di numeri a tre cifre (da 0 a 255) in notazione decimale. Ciascun numero decimale rappresenta una stringa di otto bit - se espresso in notazione binaria - e di conseguenza l'insieme dei quattro gruppi rappresenta una stringa di 32 bit.

Indirizzi IP e DNS

Ciascun utente della rete viene dunque individuato da un indirizzo IP univoco espresso da 4 numeri interi (0-255) separati dal punto «.»

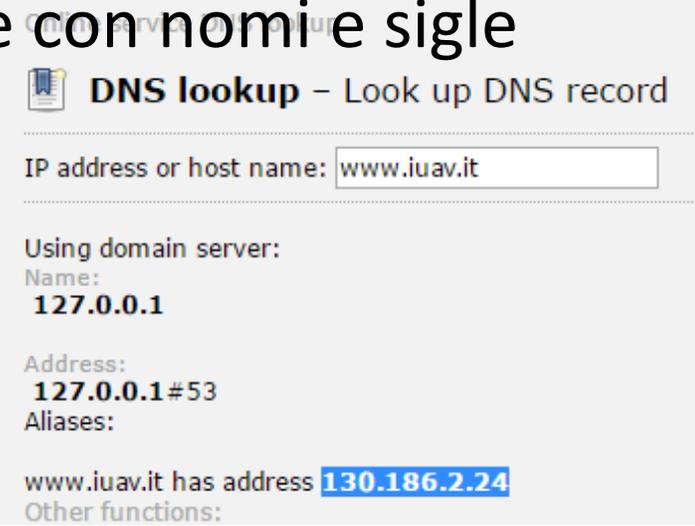
130.186.2.24

Per esempio è l'indirizzo dell'Università Iuav di Venezia.

Questo tipo di rappresentazione degli indirizzi è però molto difficile da memorizzare per cui si preferisce utilizzare un diverso sistema di rappresentazione in cui le macchine sono identificate con nomi e sigle piuttosto che con numeri:

www.iuav.it

Questo nome alternativo viene detto **dominio**.



Online Services - DNS Lookup

DNS lookup – Look up DNS record

IP address or host name:

Using domain server:
Name: **127.0.0.1**

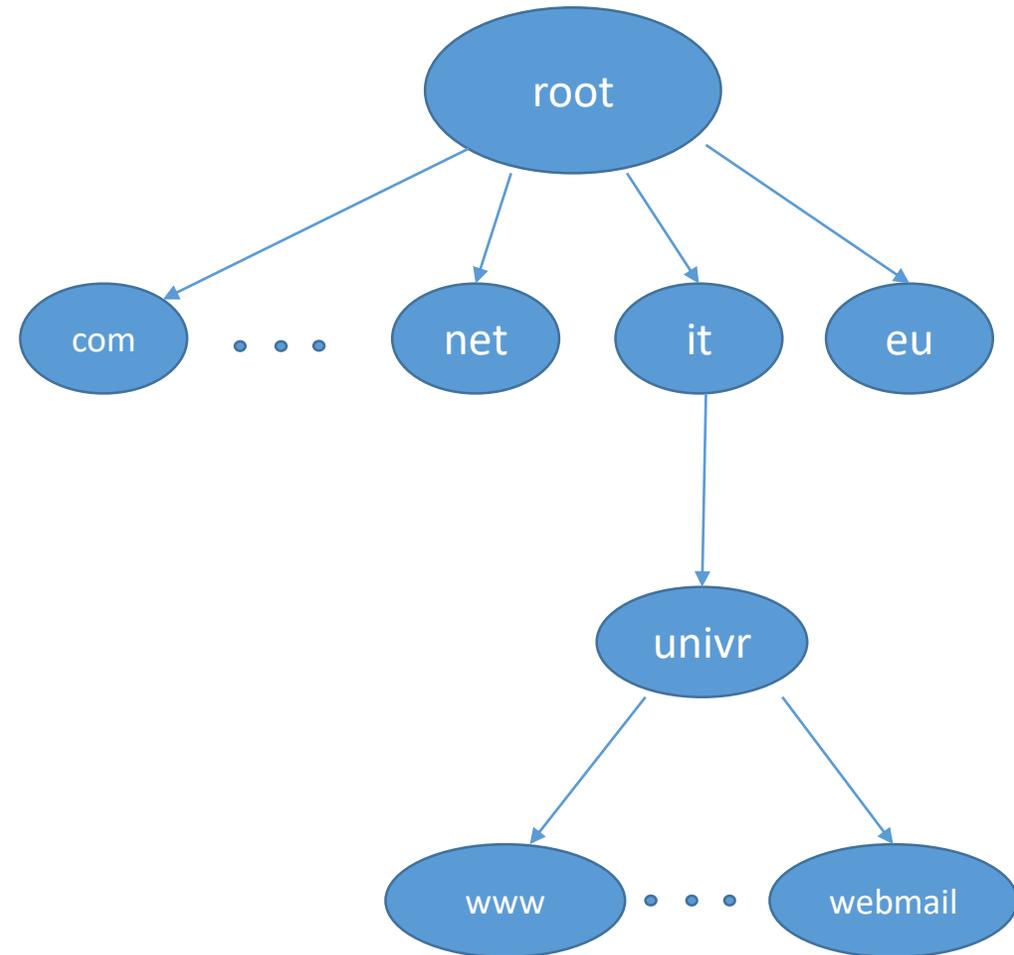
Address: **127.0.0.1#53**
Aliases:

www.iuav.it has address **130.186.2.24**
Other functions:

Indirizzi IP e DNS

Questo diverso sistema di sistema di denominazione delle risorse collegate in rete (computer, servizi, sensori etc..) fu ideata nel 1983 dall'ing. Paul Mockapretis e dall'informatico Jon Hostel.

Questo sistema viene chiamato **Domain Name System** (DNS) e da un punto di vista pratico non è altro che un DataBase Distribuito (di cui parleremo più avanti) nel quale vengono memorizzati i nomi e gli IP address di tutte le risorse Internet. La struttura del DataBase è di tipo gerarchico ed è suddivisa in «**domini**».



Indirizzi IP e DNS

Un nome di dominio è una serie di stringhe separate da punti.

Nel DNS la parte più significativa si trova nella parte destra della stringa: questa è detta **Dominio di Primo Livello** (Top Level Domain).

Sono Top Level Domain per esempio:

.it o **.uk** o **.fr** – per identificare la nazione; **.eu** per identificare l'Unione Europea; **.org** – per le organizzazioni non commerciali; **.com** – per le organizzazioni commerciali; **.edu** – per scuole e università negli USA etc..

La parte che precede il Dominio di Primo Livello si chiama: Dominio di Secondo Livello e questo normalmente si riferisce all'Organizzazione o all'Ente. Per esempio:

univr.it è un dominio di secondo livello

Indirizzi IP e DNS

- Il Dominio di secondo livello deve essere **registrato** presso un **Domain Name Register** ossia una entità commerciale che si occupa di riservare gli Internet Domain Names.
- In Italia esiste la Registration Authority Italiana «**Registro.it**» gestita dal Consiglio Nazionale delle Ricerche (CNR) che dal 1987, su delega del IANA (Internet Assigned Number Authority), si occupa di registrare i domini «.it».
- Le eventuali porzioni di dominio che precedono il secondo livello servono normalmente a specificare una particolare area o tematica sviluppata all'interno del sito: per esempio il prefisso «.en» in “en.wikipedia.org” specifica che il sito in esame si riferisce al wikipedia in lingua inglese.
- Infine il prefisso «**www**» che precede il nome del dominio «univr.it» o il «**mail**» che precede «google.com» ha lo scopo di specificare il tipo di servizio Internet che si sta utilizzando.

Il World Wide Web

A questo punto abbiamo capito, a grandi linee, come funziona Internet in quanto infrastruttura di rete, tuttavia aver parlato di protocolli, di indirizzamenti, di instradamenti, di Provider, di DNS, non ci ha fatto entrare in contatto con l'Internet che conosciamo, cioè quell'entità che ci permette di cercare informazioni, chattare con gli amici, navigare da un sito a un altro, fare acquisti on-line, guardare un film in streaming e tanto altro.

Internet infatti è una interconnessione globale di reti di computer che sono messi in grado di comunicare, tuttavia, fatta questa affermazione, è lecito chiederci:

come possiamo sfruttare questa grande opportunità?

Il World Wide Web



Per rispondere a questa domanda facciamo un salto a Ginevra nel 1989, al CERN.

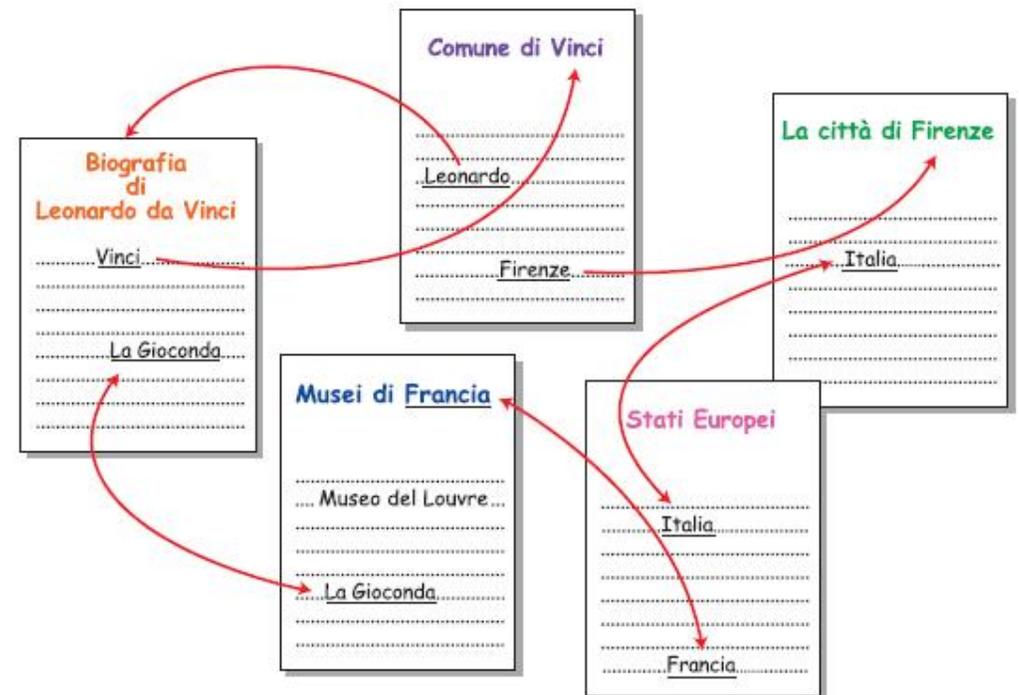
In quel periodo un ricercatore inglese, Tim-Berners Lee, proponeva un sistema di gestione delle informazioni in cui un server e un client utilizzavano l'infrastruttura Internet per comunicare attraverso un nuovo protocollo: **HTTP - Hyper Text Transfer Protocol**.

Con HTTP si entra nel mondo degli **IperTesti**, testi strutturati in grado di utilizzare dei riferimenti logici (IperLink) ad altri testi. Attraverso gli IperLink è dunque possibile spostarsi da una pagina a un'altra, recuperando informazioni distribuite sui diversi nodi della rete Internet. Questa possibilità di «navigare» nella rete è alla base del servizio più noto e più utilizzato di Internet: il World Wide Web, o WWW o ancora, semplicemente, «**the Web**».

Il World Wide Web

Sebbene nell'uso comune i termini Internet e Web siano spesso utilizzati senza distinzione, è bene precisare ancora una volta che mentre Internet è un sistema globale di reti di computer interconnessi, il Web è una interconnessione globale di documenti e di altre risorse, che ne fanno un gigantesco spazio informativo.

Un **sito web** dal canto suo non è altro che un insieme di pagine web correlate, ovvero una struttura di ipertesti che risiede su un dato server web a cui i client possono collegarsi (architettura client/server).



Il World Wide Web

Come già detto, nel WWW il protocollo standard utilizzato per la trasmissione delle informazioni si chiama **HyperText Transfer Protocol (HTTP)**. Un server HTTP rimane costantemente in «ascolto» dei client e ogniqualvolta siano richiesti i suoi servizi, utilizza il protocollo TCP per il trasporto dell'informazione. Nell'uso comune il client corrisponde al **browser** ed il server alla macchina su cui risiede il **sito web**.

Livello	Suite di protocolli Internet
Applicazioni	BGP , DHCP , DNS , FTP , HSRP , HTTP , HTTPS , IGRP , IMAP , IRC , POP3 , RIP , Rsync , RTP , RTSP , SFTP , SIP , SMTP , SNMP , SSH , Telnet , VoIP
Trasporto	DCCP , SCTP , TCP , UDP
Rete	IPv4 , IPv6 , ICMP , ICMPv6 , IGMP , IPsec , OSPF
Collegamento	Ethernet , WiFi , PPP , Token ring , ARP , ATM , FDDI , LLC , SLIP , WiMAX , HSDPA , MPLS

Il World Wide Web

Nel WWW le risorse residenti sui vari computer collegati alla rete vengono individuate con un indirizzo detto **URL (Uniform Resource Locator)** che l'utente deve esplicitare ogniqualvolta voglia raggiungere una data risorsa.

La URL contiene alcune informazioni che sono: il tipo di protocollo utilizzato (p.e. **http**) seguito dal dominio della Organizzazione (p.e. www.univr.it) e dall'indirizzo della risorsa cercata.

Per esempio **jsp/index.jsp** ci porta alla home page dell'Università

<http://www.univr.it/jsp/index.jsp>

Mentre nella URL

<http://www.dtesis.univr.it/?ent=persona&id=33426&lang=it>

Viene inoltrata al server la richiesta (dopo il «?») di recuperare i dati in lingua italiana della persona con ID 33426.

Il World Wide Web

Infine, le pagine Web che consultiamo non altro che sequenze di testi e di caratteri di controllo interpretabili dal browser e costruiti utilizzando un particolare linguaggio denominato **HTML (Hypertext Markup Language)**

Ecco un semplice esempio:

```
<H1 align="left">Questo titolo è allineato a sinistra</H1>  
<H1 align="center">Questo titolo è allineato a destra</H2>
```

in fase di presentazione nel primo caso il risultato sarà un titolo di primo livello allineato a sinistra

Questo titolo è allineato a sinistra

nel secondo caso un titolo di secondo livello allineato a destra

Questo titolo è allineato a destra

Titoli	<H1> ... </H1>
Paragrafi	<P> ... </P>
Interruzioni di riga	
Suddivisione in blocchi	<DIV> ... </DIV>
Rientri	<BLOCKQUOTE> ... </BLOCKQUOTE>
Stampa di linee orizzontali	<HR>

Le tecnologie per l'accesso alla rete

In generale un computer si dice **connesso a Internet** quando, dopo aver ottenuto l'autorizzazione del Internet Service Provider e aver ottenuto il proprio indirizzo IP (statico o dinamico), è in grado di iniziare la sua sessione di navigazione in rete e ad usufruirne dei servizi.

Tale connessione è normalmente ottenuta con un **modem** (ISDN, ADSL,..) che si connette alla linea telefonica o alla rete WiFi (chiavette usb), oppure collegandosi al **router** di una rete LAN.

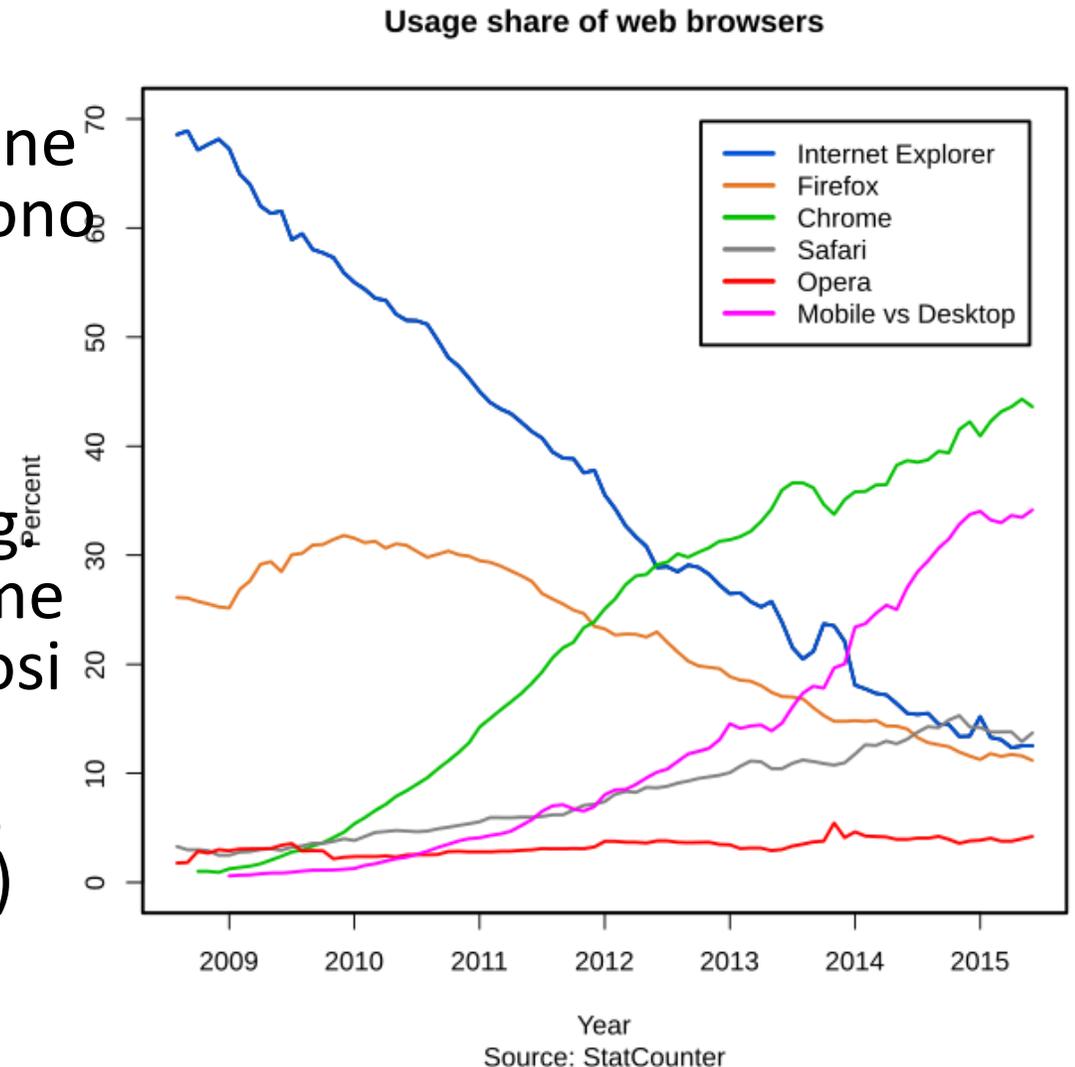
Se la connessione avviene in radiofrequenza si parla di connessioni di tipo **mobile** (3G, 4G).

Se la connessione avviene via satellite con modem dedicato e antenna parabolica si parla di connessione a banda larga **satellitare**.

Browser e motori di ricerca

Il **Browser Web** è un'applicazione per il recupero, la presentazione e la navigazione di risorse web. Tali risorse, come visto, sono messe a disposizione sotto forma di ipertesto.

Oltre a ciò i browser normalmente si occupano di gestire altri aspetti come e.g. quelli di sicurezza (la «Blacklist» di Chrome è una lista di siti potenzialmente pericolosi e l'utente viene avvisato se cerca di visitarne uno) e di aiuto alla navigazione, profilando l'utente (servizio disattivabile) per fornirgli un servizio personalizzato.



Browser e motori di ricerca

Un **Motore di Ricerca** invece è un sito web che può essere visitato utilizzando un Browser Web che.

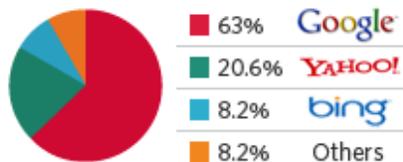
Su richiesta dell'utente il Motore di Ricerca analizza un insieme di informazioni presenti nel Web e restituisce (sempre attraverso il Browser) un indice dei contenuti disponibili.

Il processo di **indicizzazione** dei contenuti Web viene effettuato costantemente a monte di una qualsiasi richiesta, in base a criteri che variano in base al fornitore del servizio.

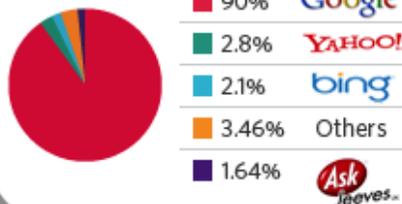
Anche la **classificazione** viene effettuata in modo automatico in base a formule statistico-matematiche che possono essere controllate per fornire servizi a pagamento.



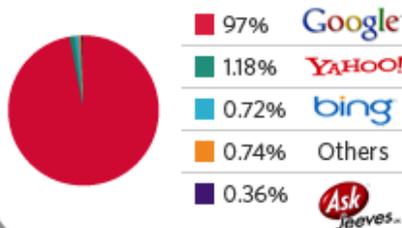
United States



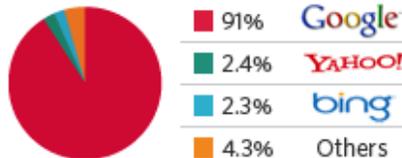
United Kingdom



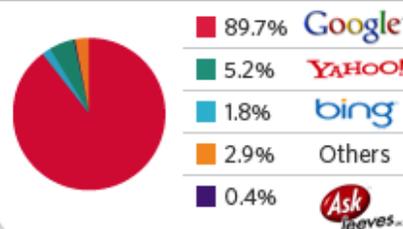
Spain



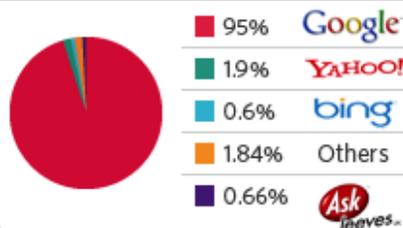
France



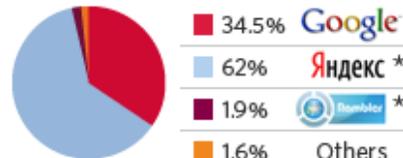
Germany



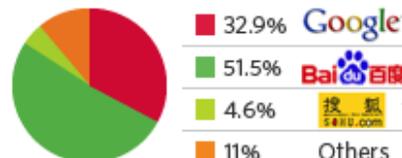
Italy



Russia



China



La Posta Elettronica

La **posta elettronica** (e-mail) è anch'esso uno dei servizi più noti e utilizzati di Internet. Il protocollo di comunicazione tra server di posta si chiama **Simple Mail Transfer Protocol** (SMTP).

In questo caso gli utenti che vogliono utilizzare il servizio debbono possedere una casella postale (**mailbox**) e delle **credenziali** di accesso su un apposito server, dove i messaggi vengono depositati.

Un tipico indirizzo di posta elettronica è composto da:

username@dominio

Le operazioni di compilazione dei messaggi possono essere effettuate **offline** mentre l'invio può avvenire in qualsiasi momento quando si è connessi alla rete e così la ricezione (**modalità asincrona**).

Il messaggio è composto da una **intestazione** (mittente, destinatari, indirizzi in cc, indirizzi in cc nascosta ccn, oggetto), da un **corpo** (il messaggio vero e proprio) e da eventuali allegati.

I Servizi di Social Networking

Un **Social Networking Service**, è una piattaforma che serve a costruire «reti e relazioni sociali» tra persone e gruppi che condividono interessi simili. Normalmente i servizi di questo tipo permettono all'utente di creare un proprio profilo, di organizzare una lista di contatti con cui comunicare, di pubblicare dei contenuti e di accedere a quelli altrui.

L'impatto dei social network negli ultimi anni è stato enorme tanto che è stato coniato il termine **Web 2.0** per definire questo nuovo fenomeno in cui le persone sono fruitori di contenuti e parte attiva nella realizzazione di nuovi contenuti.

Col termine Web 2.0 ci si riferisce all'insieme di tutte quelle applicazioni che permettono un elevato livello di interazione tra utenti come p.e. i **blog**, i **forum**, le **chat**, i **wiki** e i social network propriamente detti come **Facebook**, **Twitter**, **Google+**, **Linkedin** e molti altri.

Web 2.0

Aggregators Folksonomy Wikis User Centered Joy of Use
Blogs Participation Six Degrees Usability Widgets
Pagerank XFN Social Software FOAF Browser
Recommendation Sharing Collaboration Perpetual Beta Simplicity AJAX
Videocasting Podcasting Audio IM Video Design
Convergence CSS Pay Per Click
UMTS Mobility Atom XHTML SVG Ruby on Rails VC Trust Affiliation
OpenAPIs RSS Semantic Web Standards SEO Economy
OpenID Remixability REST Standardization The Long Tail
DataDriven Accessibility XML
Modularity SOAP Microformats Syndication

Prossimo Capitolo – Basi di Dati

Nel prossimo capitolo vedremo cosa è una Base di Dati e cosa sono un DBMS e un Data Base distribuito.