

Università Degli Studi di Ferrara

Corso di Laurea in Informatica - A.A. 2023 - 2024

Tecnologie Web

Lez. 01 - HTTP e W3C

Cap. 1

Tecnologie Server Side

In questa lezione...

- Cenni sul protocollo HTTP
- URL
- W3C

Modello ISO/OSI

Applicazione
Presentazione
Sessione
Trasporto
Rete
Dati
Fisico

L'Open Systems Interconnection (modello ISO/OSI) è uno standard formalizzato nel 1978 dal principale ente di standardizzazione internazionale, l'International Organization for Standardization (ISO) che stabilisce gli standard di comunicazione per l'architettura logica di rete.

Il modello modello ISO/OSI è un insieme di protocolli di comunicazione di rete suddiviso in 7 livelli

Modello ISO/OSI

Applicazione
Presentazione
Sessione
Trasporto
Rete
Dati
Fisico

Noi ci concentreremo sul livello più alto (Applicazione) del modello ISO/OSI all'interno del quale sono definiti i protocolli che noi utilizzeremo, primo fra tutti il protocollo HTTP, e menzioneremo un protocollo del livello di Rete (Network), il protocollo IP

Protocollo HTTP

Il protocollo HTTP (**H**yper**T**ext **T**ransfer **P**rotocol) insieme al linguaggio HTML e agli URL costituiscono il nucleo del World Wide Web (WWW)

Le specifiche di questo protocollo sono gestite dal World Wide Web Consortium (W3C <http://www.w3c.org>)

Ad oggi la versione ufficiale più recente del protocollo HTTP è la versione 3 che si indica con HTTP/3

Gli identificatori

Il protocollo HTTP fa uso degli URI (Uniform Resource Identifier) che sono una sintassi usata sul Web per identificare le **risorse** (documenti HTML, immagini, file audio, ecc.) sulla rete Internet.

Gli URI si suddividono in:

- URL: Uniform Resource Location
- URN: Uniform Resource Name

Gli URL

Gli **URL** sono particolari URI che identificano una risorsa immediatamente utilizzabile da un programma, ed hanno il seguente schema:

`<protocol>://<host>[:<port>] [<path>] [?<query>]`

Ad esempio:

<http://www.unife.it:80/informatica/logo.jpeg>

The diagram shows the URL `http://www.unife.it:80/informatica/logo.jpeg` with brackets underneath identifying its parts: `http` is labeled 'protocol', `www.unife.it` is labeled 'host name', `:80` is labeled 'port' (with an upward arrow), and `/informatica/logo.jpeg` is labeled 'path name'.

Gli URL

Gli URL contengono tutte le informazioni necessarie per accedere alla risorsa, ma sono fragili cioè possono perdere il loro “valore” se soggetti a modifiche anche non sostanziali, ad esempio il cambio del nome di una directory.

`http://www.unife.it:80/informatica/logo.jpeg`

`http://www.unife.it:80/informatica/immagini/logo.jpeg`

Gli URN

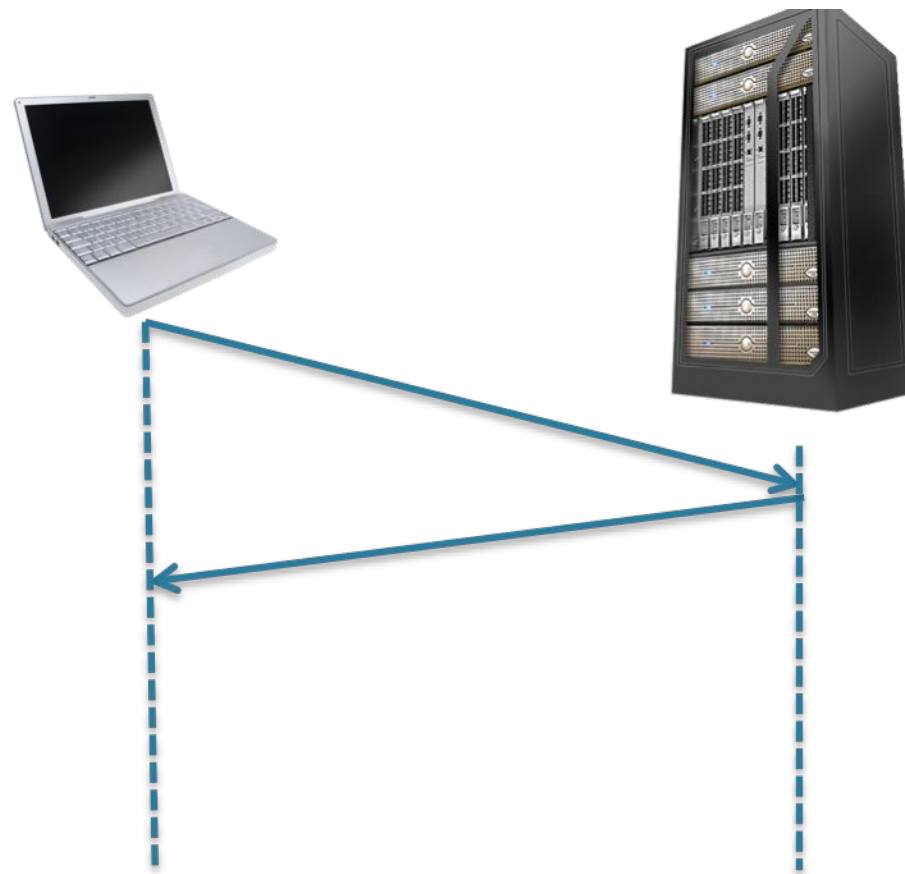
Un URN invece, è un nome stabile e definitivo di una risorsa, che fornisce un'informazione certa e affidabile sulla sua esistenza e accessibilità.

Gli URN devono essere trasformati da un apposito servizio, negli URL attualmente associati alle risorse. La mappa quindi, deve essere aggiornata ogni volta che la risorsa viene spostata.

HTTP Esempio

1) Il client invia una richiesta al server attraverso un URL

3) Il client riceve il contenuto



2) Il server risponde

Protocollo HTTP

Il protocollo funziona su un meccanismo client/server in cui due entità (il client ed il server) si scambiano messaggi di richiesta e di risposta.

Il messaggio di richiesta che viene inviato dal client al server è composto da 3 parti:

Request
Header
Body

HTTP Request

La riga di richiesta è composta da: metodo (o verbo), URI e versione del protocollo, ad esempio:

```
GET /uri/risorsa HTTP/1.1
```

I metodi previsti per la versione 1.1 del protocollo sono:

```
GET, POST, HEAD, PUT, DELETE, TRACE, OPTIONS, CONNECT
```

Sono generalmente indicati con lettere maiuscole ed i primi 5 sono quelli più utilizzati.

HTTP Header

Nell' header della richiesta HTTP vengono inserite delle indicazioni sul formato della richiesta atteso ed altre informazioni, ad esempio:

- User-Agent:
- ContentType:
- Accept:
- Authentication:

Queste informazioni vengono inserite come elenco di:

Chiave: Valore

HTTP Body

Il corpo del messaggio contiene le informazioni che vogliamo trasferire al server.

Protocollo HTTP

Il messaggio di risposta che viene inviato dal server al client è anch'esso composto da 3 parti:

Header
Status
Body

HTTP Header

Come per la richiesta nell' header della risposta vengono inserite delle indicazioni sul formato dei dati inviati altre informazioni, ad esempio:

- Content-Type:

Anche queste informazioni vengono inserite come elenco di:

Chiave: Valore

HTTP Status

Lo Status che viene inviato dal server al client è composto da una riga contenente un codice a 3 cifre suddivisi in “categorie”. Ogni categoria inizia con un codice specifico e determina cosa contiene la risposta:

- 1xx: Messaggi informativi
- 2xx: OK (richiesta soddisfatta, nel body ci sarà il contenuto)
- 3xx: Redirezione ad un'altra risorsa
- 4xx: Errori nella richiesta
- 5xx: Errore da parte del server

HTTP Body

Il corpo del messaggio contiene le informazioni che abbiamo richiesto al server.

Esempio Richiesta

Request	GET /unife.it/index.html HTTP/1.1
Header	User-Agent: Mozilla/5.0 (compatible; Konqueror/3.2; Linux) (KHTML, like Gecko) Accept: text/html, image/jpeg, image/png, text/*, image/*, */* Accept-Encoding: x-gzip, x-deflate, gzip, deflate, identity Accept-Charset: iso-8859-1, utf-8;q=0.5, */q=0.5 Accept-Language: it
Body	id=1234

Esempio Risposta

Status	HTTP/1.0 200 OK
Header	Date: Mon, 28 Jun 2016 10:47:31 GMT Server: Apache/1.3.29 (Unix) PHP/4.3.4 X-Powered-By: PHP/4.3.4 Content-Language: it Content-Type: text/html; charset=utf-8
Body	<!DOCTYPE html> <html xmlns="http://www.w3.org/1999/xhtml" lang="it"> ..

HTTP Limiti

Il protocollo HTTP ed il linguaggio HTML sono stati originariamente progettati con lo scopo di visualizzare e navigare ipertesti.

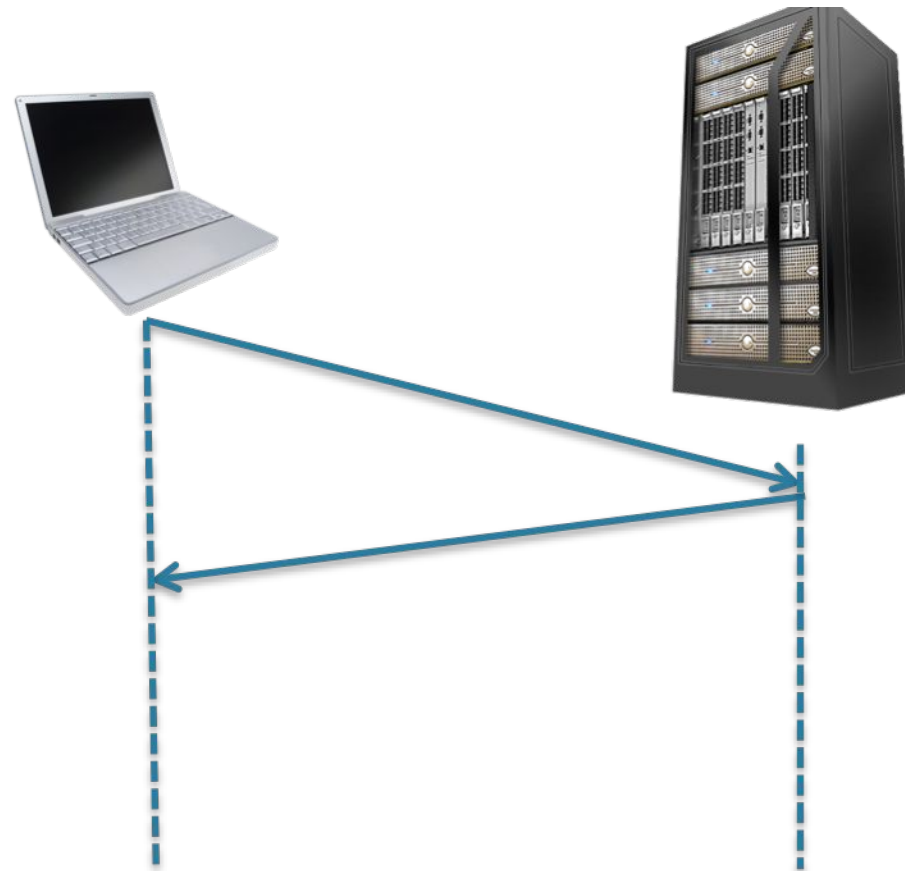
In particolare il protocollo HTTP:

- E' **stateless**: Ogni richiesta è indipendente da quella precedente;
- E' una tecnologia pull: Per ottenere informazioni è necessario che il client ne faccia esplicita richiesta;
- Non descrive l'origine di una richiesta: (click o visita diretta)

HTTP Esempio

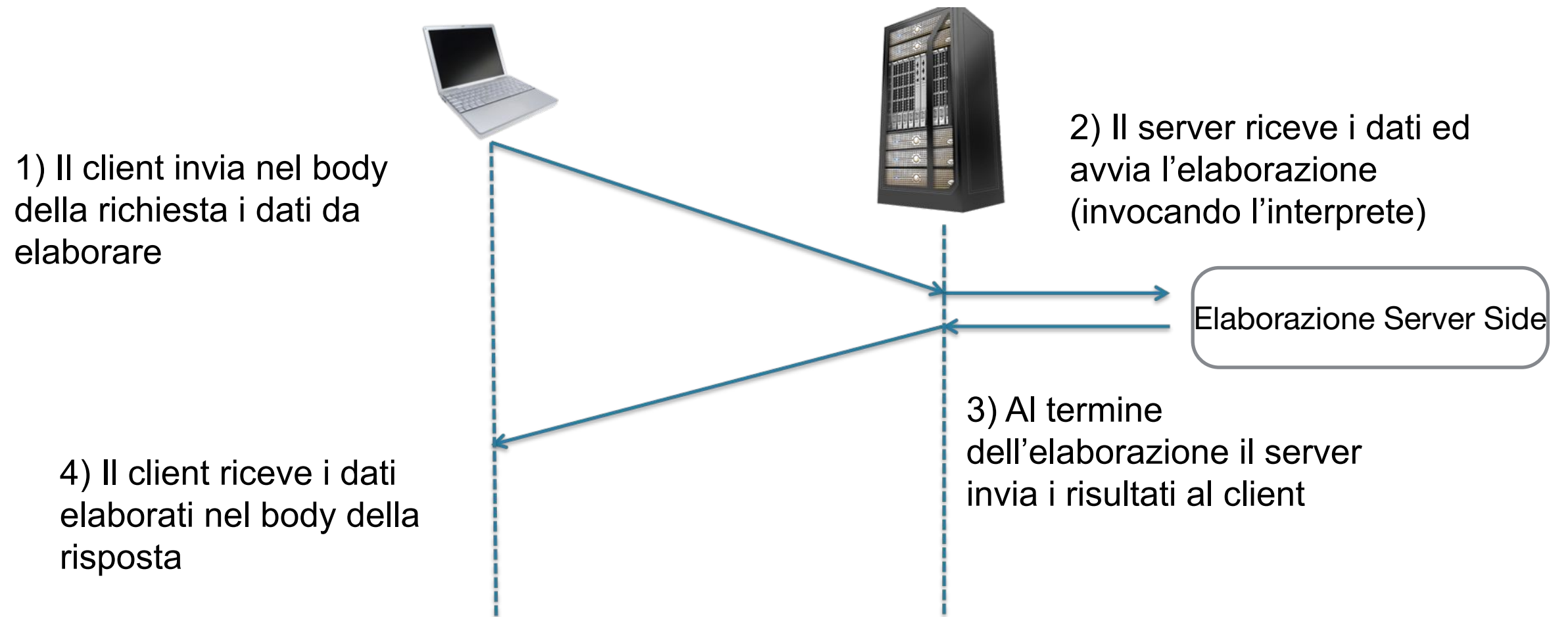
1) Il client richiede una risorsa al web server attraverso un URL

3) Il client riceve la risorsa nel body della risposta



2) Il server risponde includendo nella risposta il contenuto della risorsa

HTTP Esempio (2)

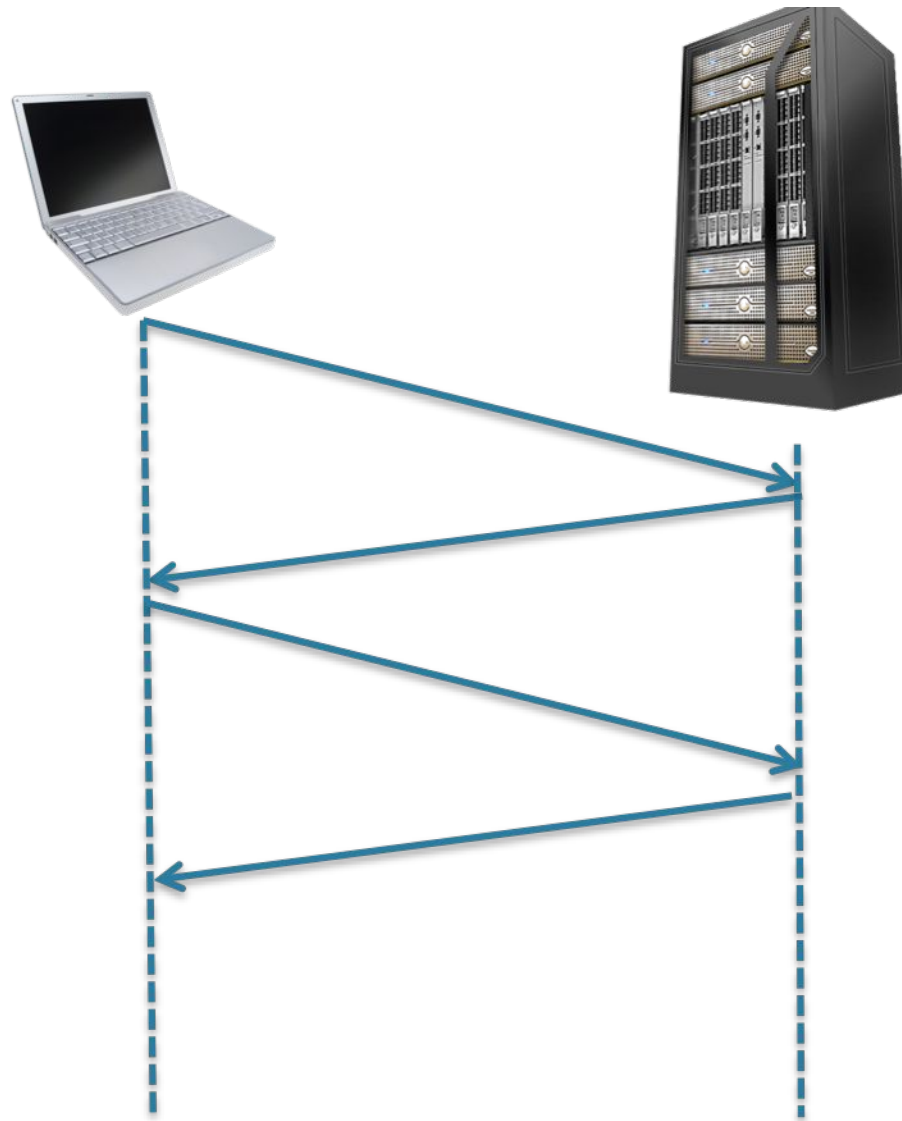


HTTP Esempio Autenticazione

1) Il client invia la richiesta per ottenere una risorsa

3) Il client richiede all'utente username e password e ripete la richiesta includendo le credenziali nell'header nella chiave "Authorization"

5) Il client riceve la risorsa nel body della risposta



2) Il server risponde con una risposta 401 (non autorizzato)

4) Il server verifica le credenziali ed invia la risorsa

HTTP/2

Nel Maggio del 2015 è stato standardizzata (da IETF con la RFC 7549) la nuova versione del protocollo HTTP (HTTP/2) e oggi è supportato da tutti i browser.

La standardizzazione del protocollo HTTP/1.1 risale al 1997.

Non è una ristrutturazione del protocollo HTTP/1.1 e garantisce retrocompatibilità lasciando inalterati gli header, status code e metodi. È un 'wrap' del protocollo HTTP/1.1 sul protocollo SPDY di Google che riduce i tempi di latenza del protocollo HTTP/1.1

Maggiori dettagli su: <http://dev.chromium.org/spdy>

HTTP/2 Performance

Multiplexing: Nella stessa connessione è possibile avere richieste e risposte multiple che vengono “ricostruite” nell’host di destinazione. Questo permette di fare operazioni in parallelo e quindi di ridurre la latenza.



HTTP/2 Performance

Compression headers: Anche gli header HTTP sono compressi e il consumo di banda è notevolmente ridotto.

Supporto per operazioni di **Push** da parte del server, che può inviare dati al client prima che questo ne faccia richiesta.

Il Protocollo IP

Il protocollo IP appartiene al gruppo dei protocolli di Rete.
Sul protocollo IP si basa l'identificazione dei nodi della rete, ad ogni nodo deve essere associato un indirizzo del protocollo IP che ha la seguente forma: XXX.XXX.XXX.XXX

Ogni indirizzo IP è formato da 32 bit suddivisi in 4 gruppi da 1 Byte

Ogni byte può avere un valore da 0 a 255 e quindi gli indirizzi IP possono variare da 0.0.0.0 a 255.255.255.255

Il Protocollo IP

Noi non ci soffermeremo sul questo protocollo ma un paio di cose è necessario sapere:

Esiste un sistema per risolvere (convertire) i nomi di rete (identificatori) negli indirizzi IP e viceversa. Questo sistema si chiama DNS.

Esistono degli indirizzi IP “riservati”, cioè indirizzi ai quali non può essere associato nessun dispositivo di rete perché per convenzione sono già utilizzati per altro, il più famoso è 127.0.0.1 che identifica sempre la macchina locale

W3C

Il World Wide Web Consortium (W3C) viene fondato nel 1994 da T. Berners-Lee (che tuttora ne è il direttore) come collaborazione tra il MIT (Boston) e il CERN (Ginevra), con lo scopo di:

"Condurre il World Wide Web al suo pieno potenziale sviluppando protocolli e linee guida che assicurino una crescita di lungo periodo del Web".

W3C

In pratica il W3C è un consorzio internazionale in cui le organizzazioni membro, lo staff permanente del consorzio ed il pubblico lavorano insieme per sviluppare gli standard del Web.

Ciclo di vita degli standard

Ogni standard, prima di diventare tale, cioè definitivo, subisce diversi passaggi evolutivi della sua definizione:

- Working Draft (WD);
- Candidate Recommendation (CR);
- Proposed Recommendation (PR);
- W3C Recommendation (REC)

W3C - WD

Working Draft

Un documento allo stadio di Working Draft è un documento che è stato rilasciato dal W3C per essere rivisto dalla comunità.

W3C - CR

Candidate Recommendation

È un documento che ha subito ampie revisioni e soddisfa i requisiti tecnici del gruppo di lavoro.

Viene rilasciata per avviare la sperimentazione delle implementazioni.

W3C - PR

Proposed Recommendation

Documento maturo, con ampia revisione e verificate possibilità di implementazione in attesa dell'approvazione finale.

W3C - REC

Recommendation

Documento di fatto, uno standard per il W3C che ne raccomanda il massimo utilizzo e diffusione.

La situazione attuale

Navigando su <https://www.w3.org/TR/> e usando i filtri riempite idealmente questa tabella inserendo 2 o 3 righe.

Quale è la tecnologia più interessante che avete trovato? (rispondete su Classroom)

#	WD	CR	PR	REC
1				
2				
3				

Domande?