



Hyper Text Markup Language - HTML

Introduzione e funzionamento dei Browsers

Sviluppo di siti web – UD01

prof. Giovanni Borga

L'HTML e i browser

L'HTML (Hyper Text Markup Language) è il linguaggio con cui si indica come i vari elementi vanno disposti in una pagina Web.

Un documento html non è nient'altro che un **file di testo** con delle indicazioni sul colore delle scritte, sulla posizione delle immagini all'interno della pagina, su come far scorrere il testo, e altre cose di questo genere.

L'HTML NON è un linguaggio di programmazione. Più propriamente è un linguaggio di «formattazione»

Il Browser è il programma che si usa per navigare nel Web. Svolge principalmente due compiti:

- scarica i vari files che si trovano su un computer remoto (il server) e che fanno riferimento a un determinato indirizzo
- legge i documenti scritti in html e, a seconda delle indicazioni ivi contenute, visualizza la pagina in un modo, piuttosto che in un altro; inoltre i vari files associati a quel documento (ad esempio le immagini, o i contenuti multimediali) vengono disposti secondo le indicazioni del codice

I browser attualmente più diffusi:
Chrome, Firefox, Internet Explorer / Edge,
Safari, Opera



L'HTML e i browser

Acquisire una mentalità multi-browser

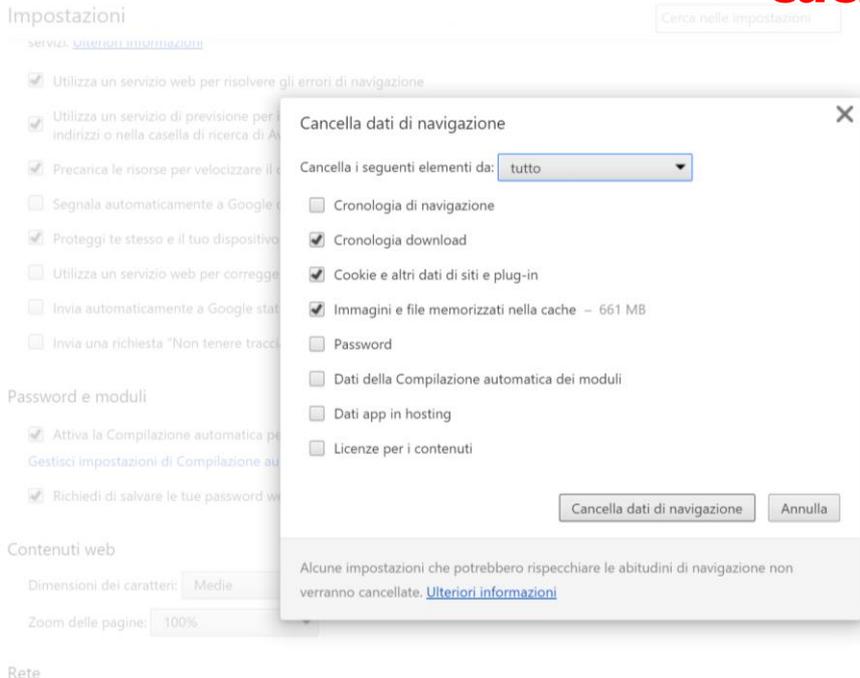
Per lo sviluppo dei siti non occorre tanto conoscere nei minimi dettagli il codice HTML, quanto piuttosto sapere come il codice HTML verrà visualizzato sul computer dell'utente.

Infatti uno dei lavori più difficili è quello di riuscire a far vedere correttamente il proprio sito con i browser e le piattaforme più svariate.

La cache

I files scaricati dal web vengono memorizzati in una particolare cartella del computer che prende il nome di:

cache



Può succedere - soprattutto a chi è alle prime armi - di continuare a modificare un file, ma di non riuscire a vederne le modifiche. Questo succede perché la pagina visualizzata è sempre quella vecchia memorizzata nella cache.

In molti browser è possibile impostare le preferenze in modo che venga sempre cancellata la cache quanto si chiude il programma.

Motore di rendering

La visualizzazione di un file html da parte del browser prende il nome di

rendering della pagina

Il cosiddetto «**motore di rendering**» è quella sezione del browser che si occupa di visualizzare la pagina applicando stili, layout ed elementi *embedded* secondo quanto indicato nel codice HTML.

In qualsiasi momento è possibile visualizzare il codice HTML delle pagine che stiamo visitando utilizzando la funzione di visualizzazione sorgente:



Funzionamento di un browser

Hypertext Markup Language ("Linguaggio di contrassegno per gli Iper testi") non è un linguaggio di programmazione ma un linguaggio di contrassegno (o 'di marcatura'), che permette di indicare come disporre gli elementi all'interno di una pagina: le indicazioni vengono date attraverso degli appositi marcatori, detti "**tag**".

- *Alcuni linguaggi di programmazione sono: C, C++, Visual Basic, Pascal, Java;*
- *Alcuni linguaggi di scripting sono: PHP, ASP, PERL, JavaScript.*

Ciò significa che l'HTML:

- a) non ha meccanismi che consentono di prendere delle decisioni ("in questa situazione fai questo, in quest'altra fai quest'altro")
- b) Non è in grado di compiere delle iterazioni ("ripeti questa cosa, finché non succede questo"), né ha altri costrutti propri della programmazione.

Pur essendo dotato di una sua sintassi, non presuppone la logica ferrea e inappuntabile dei linguaggi di programmazione: se vi dimenticate di chiudere un tag, non verranno prodotti dei messaggi di errore; se non rispettate la sintassi probabilmente non otterrete la visualizzazione della pagina che desiderate, ma nient'altro. A volte può essere necessario adottare dei "trucchetti", non proprio da manuale, pur di visualizzare la pagina correttamente con i diversi browser.

Lo standard HTML

L'organizzazione che si occupa di standardizzare la sintassi del linguaggio HTML è il **W3C: World Wide Web Consortium**.

Ha rilasciato diverse versioni di questo linguaggio (HTML 2.0, HTML 3.2, HTML 4.0); e - da un certo punto in poi - l'HTML si è evoluto in **XHTML** (HTML riformulato con le regole sintattiche del XML).

La versione dell'HTML che utilizzeremo in questo corso è la 4.01 rilasciata il 24 dicembre 1999.

In realtà quasi tutti i browsers supportano già da tempo la **nuova versione HTML 5** che però non è rilasciata ufficialmente come release ma come **recommendation**.

HTML 5 non soppianta la versione 4 ma integra nuove funzionalità e ottimizza alcuni aspetti.



HTML 4.01 Specification

W3C Recommendation 24 December 1999

This version:

<http://www.w3.org/TR/1999/REC-html401-19991224>
([plain text \[794Kb\]](#), [gzip'ed tar archive of HTML files \[371Kb\]](#), [a \[963Kb\]](#))

Latest version of HTML 4.01:

<http://www.w3.org/TR/html401>

Latest version of HTML 4:

<http://www.w3.org/TR/html4>

Latest version of HTML:

<http://www.w3.org/TR/html>

Previous version of HTML 4.01:

<http://www.w3.org/TR/1999/PR-html40-19990824>

Previous HTML 4 Recommendation:

<http://www.w3.org/TR/1998/REC-html40-19980424>

Editors:

[Dave Raggett](mailto:dsr@w3.org) <dsr@w3.org>

Arnaud Le Hors, W3C

Ian Jacobs, W3C

HTML 5

Principali innovazioni

- Regole più stringenti per la **strutturazione del testo** in capitoli, paragrafi e sezioni;
- Nuovi elementi di controllo per i **menu di navigazione**;
- Miglioramenti ed estensione degli elementi di controllo per i **moduli**;
- Nuovi elementi specifici per il controllo di **contenuti multimediali** (tag <video> e <audio>);
- **Estensione degli attributi** di quasi tutti i tag, specialmente quelli finalizzati all'accessibilità, finora previsti solo per alcuni tag;
- Supporto per il **Canvas** che permette di utilizzare JavaScript per creare animazioni e grafica bitmap;
- Introduzione della **geolocalizzazione**, dovuta ad una forte espansione di sistemi operativi mobili (quali Android e iOS, tra i più diffusi);
- Introduzione di un sistema alternativo ai normali cookie, più efficiente, chiamato **Web Storage** che consente anche un notevole risparmio di banda;
- Standardizzazione di programmi JavaScript, chiamati **Web Workers** e possibilità di utilizzare siti offline;
- Sostituzione del tag **doctype** inutilmente prolisso, con un semplice <!DOCTYPE html>

Deprecati o eliminati elementi di scarso o nessun utilizzo effettivo

Estensioni dei files

Apriete o create un documento con il blocco note, e salvate il file in qualche cartella del vostro computer con **estensione "html"**, ad esempio **miaPagina.html**.

Una pagina web è «semplicemente» un file ASCII con estensione .html

Fino a qualche tempo fa si era soliti attribuire ai file l'estensione htm, ma questo avveniva perché il dos e poi Windows 3.1 non erano in grado di gestire i file con nomi di grandezza superiore a 8 caratteri ed estensione superiore alle 3 lettere. Dunque .html era diventato .htm, così come .jpeg era diventato .jpg.

Si possono utilizzare liberamente entrambe le estensioni ma ormai la consuetudine è quella di utilizzare .html

Normalmente il sistema operativo associa automaticamente le estensioni .htm e .html al browser predefinito.

NB: Se accade che il browser non «riconosce» la vostra pagina html verificate di non aver salvato con una «doppia estensione» (es. .html.txt). Questo può succedere se il vostro sistema operativo è impostato per nascondere le estensioni per i files già associati ad un'applicazione.

