

CONCETTI PRELIMINARI

- L'**HTML** (acronimo di *HyperText Markup Language* che tradotto sta per *linguaggio di marcatura¹ per ipertesti²*) è un linguaggio usato per descrivere la struttura dei documenti ipertestuali disponibili nel World Wide Web. Tutti i siti web sono scritti in HTML, codice che viene letto ed elaborato dal browser, il quale genera la pagina che viene visualizzata sullo schermo del computer.

L'HTML non è un linguaggio di programmazione, ma un linguaggio di markup, ossia descrive il contenuto, testuale e non, di una pagina web. Punto HTML (.html) o punto HTM (.htm) è anche l'estensione comune dei documenti HTML.

Questo linguaggio è stato sviluppato alla fine degli anni '80 da Tim Berners-Lee al CERN di Ginevra. Verso il 1994 ha avuto una forte diffusione, in seguito ai primi utilizzi commerciali del web.

La sintassi dell'HTML è stabilita dal World Wide Web Consortium (W3C). Durante gli anni l'HTML ha subito molte revisioni e miglioramenti, che sono stati indicati secondo la classica numerazione usata per descrivere le versioni dei software. Attualmente l'ultima versione disponibile è la versione 4.01, resa pubblica il 24 dicembre 1999.

Il componente principale della sintassi di questo linguaggio è l'*elemento* (tag), inteso come struttura di base a cui è delegata la funzione di formattare i dati o indicare al browser delle informazioni. Ogni elemento (tag) è costituito da una sequenza di caratteri racchiusi da due parentesi angolari.

- Quando si diffuse l'uso di inviare pagine web a dispositivi (ad esempio smartphone) dotati di risorse hardware e software non sufficienti ad interpretare il linguaggio HTML allora ci si accorse che era necessario un linguaggio dotato di una sintassi meglio definita: l'**XHTML** (acronimo di *eXtensible HyperText Markup Language*, Linguaggio di marcatura di ipertesti estensibile)³ che prevede un uso più restrittivo dei tag HTML.

L'XHTML, nato ufficialmente il 26 gennaio 2000, è un linguaggio di marcatura che permette di descrivere il contenuto di una pagina web in modo strutturato demandando la formattazione e la disposizione dei contenuti ai fogli di stile CSS.

La maggior parte dei browser di oggi sono in grado di interpretare correttamente i documenti XHTML e per fare ciò essi si avvalgono anche di una dichiarazione (DTD = Document Type Definition) che deve essere sempre presente in un documento XHTML.

Abbiamo detto che un documento XHTML deve necessariamente avere un DTD. Ne esistono di tre tipi, dipende dal grado di ristrettezza e/o della struttura che si intende realizzare. Vediamole qui di seguito:

- `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">`
Specifica l'uso più rigoroso di XHTML, il documento XHTML scritto utilizzando questa doctype non deve fare uso di alcun tag e/o attributo relativo all'aspetto grafico del documento stesso (font, colori, ecc.), l'aspetto grafico è lasciato esclusivamente ai fogli di stile (CSS).
- `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">`
Ammette l'utilizzo dei tags e/o attributi relativi all'aspetto grafico del documento XHTML.
- `<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">`
Prevede l'utilizzo dei frames (suddivisione della finestra visualizzata dal browser in diverse sottofinestre) nei documenti XHTML.

Per le nostre lezioni adopereremo un DTD Transitional ed un foglio di stile esterno.

Un documento XHTML che risponde alle specifiche dettate dal W3C è definito *documento valido*. La validazione (cioè il *controllo sintattico*) dei documenti XHTML non assicura al 100% la compatibilità con tutti i browser, ma è comunque fortemente consigliata. Uno strumento per controllare la validità dei documenti è disponibile presso il *Markup Validation Service* del W3C.

¹ Il termine *markup* (o *marcatura*) deriva dall'ambiente tipografico dove si usava marcare con annotazioni le parti del testo che andavano evidenziate o corrette, allo scopo di segnalarle al compositore o al dattilografo.

² Un *ipertesto* è un insieme di documenti messi in relazione tra loro tramite parole chiave. È possibile leggere, all'interno di un ipertesto, tutti i documenti collegati dalla medesima parola chiave. Il sistema delle parole chiave permette di passare da un documento all'altro rendendo possibile una lettura non sequenziale dei contenuti.

³ Va tenuto presente che più è generica la sintassi di un linguaggio, maggiore è la difficoltà di realizzare dispositivi in grado di interpretarlo correttamente.