

IL WEB CONCETTI

Termini chiavi, web publishing

parte seconda

INTRODUZIONE AD HTML5

- Gli elementi di HTML5 si dividono in tre categorie
 - Normali
 - Vuoti
 - Provenienti da altri namespaces

ELEMENTI VUOTI

- **2. Elementi vuoti:** gli elementi vuoti sono quelli che non possono avere alcun contenuto. Per questi elementi si utilizza un tag 'vuoto'. Essi sono: area, base, br, col, command, embed, hr, img, input, keygen, link, meta, param, source, track, wbr.
 - Per gli elementi vuoti, la chiusura del tag, **obbligatoria** in XHTML, è invece **opzionale**. Possiamo dunque definire un tag `` secondo le regole XHTML:
 - ``
 - OPPURE
 - ``

PROVENIENTI DA ALTRI NAMESPACES

- **3. Elementi provenienti da altri namespace:** per questi elementi sono richiesti i tag ‘autochiudenti’. Si tratta degli elementi adottati da specifiche esterne, come SVG e MathML.

ELEMENTI NORMALI

- **1. Elementi normali:** sono quelli che possono racchiudere dei contenuti sotto forma di testo, commenti HTML, altri elementi HTML, etc. Sono elementi normali, dunque, i paragrafi (`<p>`), le liste (``), i titoli (`<h1>`), etc. Salvo specifici casi, cui accenneremo nel seguito della lezione, gli elementi normali vengono definiti attraverso **un tag di apertura** (`<p>`) e **un tag di chiusura** (`</p>`).

DOCTYPE

- Poiché, come visto, le attuali specifiche di HTML lo definiscono come estensione specifica di XML, occorre definire il tipo del documento
- Questa operazione si fa specificando un riferimento ad un DTD, ovvero un Document Type Definition
- Si tratta di un file che descrive la sintassi del linguaggio impiegato, ovvero di HTML

DOCTYPE

- Non è di fatto sempre obbligatorio specificare il DocType di una pagina; la maggior parte dei browser la visualizzeranno correttamente anche senza specifica.
- Il DocType, nella sua forma più estesa, ha bisogno di diverse informazioni:
 - una dichiarazione in cui si determina se il documento è di dominio pubblico o meno;
 - una nota che veicoli a quali specifiche del W3C si sta facendo riferimento;
 - la lingua del documento.

SPECIFICHE CORRENTI

- La maggior parte delle pagine web, attualmente, utilizza questo DocType:
 - `<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//IT" http://www.w3.org/TR/html4/loose.dtd>` Il linguaggio usato all'interno della pagina è l'html, il documento è di dominio pubblico e rispetta le specifiche del W3C del tipo HTML Transitional definiti nel DTD (Document Type Definition) loose.dtd, la lingua impostata è l'italiano.

SPECIFICHE TIPICHE

- **Frameset:** questa DTD è utilizzata dai siti che al loro interno contengono i frames.
- **Strict:** è una DTD che applica rigide regole ed applica quindi i nuovissimi standard del web. Ad esempio i tag deprecati non sono ammessi e gli elementi grafici di una pagina web devono essere definiti tramite i fogli di stile.
- **Transitional:** come è facilmente intuibile dal nome questo è documento che contiene delle specifiche di transizione da uno standard all'altro, è il tipo maggiormente utilizzato ora sul web. Il documento non applica rigide regole ed è quindi possibile utilizzare tag deprecati.

ESEMPI DI SPECIFICHE

- Esempio di DocType per un sito con i frames:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"  
http://www.w3.org/TR/html4/frameset.dtd>
```

- Esempio di DocType per un sito che applica rigidamente le ultime direttive del W3C:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Strict//EN"  
http://www.w3.org/TR/html4/strict.dtd>
```

INTESTAZIONE

- All'interno del tag `<head>` possono essere inseriti, oltre al titolo, altri tag:
 - I meta tag generici
 - I meta tag `http-equiv`

META TAG GENERICI

Nome	Valore	Funzione
DC.Title	Titolo del documento	Serve a definire il titolo del documento
description	Descrizione del sito	Serve per descrivere il contenuto del sito o della pagina
creation_date	Data	Indica la data di creazione della pagina o quando essa verrà aggiornata
keywords	Ogni parola chiave separata da	Serve a dichiarare le parole chiave del sito o della pagina
robots	PROSSIMO LUCIDO	Serve a dare indicazione ai motori di ricerca per l'indicizzazione
revisit-after	n° di giorni seguito da days	Serve a dire al motore di ricerca dopo quanti giorni dovrà rivisitare la pagina
generator	L'editor usato per scrivere il codice	Indica il nome dell'editor che si è usato per scrivere il codice della pagina
copyright	Il proprietario del contenuto del sito	Indica a chi spetta il copyright del contenuto del sito
author	L'autore della pagina	Indica l'autore della pagina ed eventualmente il suo indirizzo e-mail
owner	Proprietario del sito	Indica il proprietario del sito
language	Una o più lingue	Indica la lingua o le lingue del sito
DC.Language	Una o più lingue	Indica al motore di ricerca di indicizzare le pagine in base alla lingua

IL META TAG ROBOTS

- Il meta tag *robots* serve a dare alcune indicazioni ai programmi dei motori di ricerca (i cosiddetti *spider*); più nello specifico, serve a comunicare allo spider del motore di ricerca come si deve comportare
- I suoi possibili valori sono:
 - *none*: non viene indicizzata la pagina e tutto ciò che è al suo interno
 - *index*: indicizza solo la pagina che il motore di ricerca sta visitando
 - *noindex*: non indicizza la pagina che il motore di ricerca sta visitando
 - *follow*: non indicizza la pagina che sta visitando prosegue attraverso i link della pagina
 - *nofollow*: indicizza la pagina che sta visitando ma salta i link della pagina
 - *all*: è l'insieme tra follow e index, indicizza la pagina che sta visitando e prosegue attraverso i links.

HTTP-EQUIV

La loro sintassi, in linea generale, è la seguente: `<meta http-equiv="nome_meta_tag" content="valore_del_meta_tag">`.

Nome Tag	Valore	Funzione
expires	Ora nel formato GMT	Indica la data di scadenza della validità della pagina.
reply-to	Indirizzo mail	Definisce un indirizzo e-mail a cui gli utenti possono fare riferimento
Set-Cookie	PROSSIMO LUCIDO	Salva un cookie sul computer del visitatore
refresh	PROSSIMO LUCIDO	Serve per ricaricare la pagina dopo un determinato tempo
content-Type	PROSSIMO LUCIDO	Indica il set di caratteri in uso nella pagina
Pragma	no-cache, costringe il browser a svuotare la cache e ricaricarla	Solo per NetScape. Forza il browser a non leggere il sito della cache
imagetoolbar	yes abilita la visualizzazione della toolbar, no la disabilita	Da IE 6.0 in su, abilita/disabilita la toolbar che appare sulle immagini.
distribution	global se è un contenuto di interesse generale, in caso contrario, usare local	Indica se il contenuto della pagina è di interesse generale o specifico

I COOKIE

- Il meta tag *Set-Cookie* serve, come suggerisce il suo nome, a salvare un cookie sul computer di chi sta visitando la pagina. La sua sintassi è la seguente: `<meta http-equiv="Set-Cookie" content="cookievalue=nome cookie; expires=data di scadenza; path=percorso nella cache">`
 - **cookievalue**: attraverso questo parametro impostiamo il nome che prenderà il cookie sul computer dell'utente.
 - **expires**: impostando una data in formato GMT si può determinare la data dopo la quale il cookie non avrà più effetto.
 - **path**: dove verrà salvato all'interno della cache

ESEMPIO DI SETTAGGIO DI UN COOKIE

- Esempio di settaggio di un cookie:

```
<head>
```

```
  <meta http-equiv="Set-Cookie"
```

```
    content="cookievalue=xxx;
```

```
      expires=Tuesday, 13-Jul-06 23:00:00 GMT;
```

```
      path="/">
```

```
</head>
```


REFRESH

- Il meta tag *refresh* ricarica il contenuto della pagina dopo un determinato intervallo di tempo, si può aggiungere anche un parametro mediante il quale il ricaricamento porterà ad un'altra pagina (reindirizzamento). La sua sintassi è: `<meta http-equiv="REFRESH" content="intervallo di tempo per ricaricare; url=pagina a cui porterà il ricaricamento, opzionale">`

ESEMPIO

- Esempio di reindirizzamento:

```
<meta http-equiv="refresh"  
    content="4;  
    url=http://it.wikibooks.org">
```

CONTENT-TYPE

- Il meta tag *content-Type* serve per dichiarare il set di caratteri usati all'interno della pagina. La sua sintassi è la seguente: `<meta http-equiv="content-Type" content="tipo di documento; charset=set di caratteri">`. Il tipo di documento nel nostro caso è *text/html* con l'avvento dell'xml potrebbe esserci qualche nuovo valore.

ESEMPIO

- Esempio di come utilizzare il meta tag *content-Type*:

```
<meta http-equiv="content-Type"  
      content="text/html;  
            charset=iso-8859-1">
```

RIFERIMENTI

- <http://it.wikibooks.org/wiki/HTML>