

Copyright (c) 2008, 2009, 2010 Fabio Proietti

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Authors and contributors  
Fabio Proietti

Feedback  
Please direct any comments or suggestions about this document to  
fabio.proietti AT istruzione DOT it

Publication date and version  
2008-11-15, v.0.4  
2009-11-15, v.0.5  
2010-04-21, v.0.6

Mozilla Firefox e Microsoft Internet Explorer sono nomi registrati dai rispettivi proprietari.

last modified 2012-12-01

## I commenti

Nel codice della pagina web possono essere inseriti dei commenti utili, in futuro, per chi dovrà modificare la pagina. I commenti chiariscono il significato del codice HTML

```
<!-- scrivo quello che voglio, il browser ignorerà questa riga di commento -->
```

## DTD

DTD significa "doctype declaration" (dichiarazione di tipo di documento). quindi non è un vero "tag" come <p>, ma solo una dichiarazione utile al browser. Questa dichiarazione si deve trovare sulla prima riga del codice della pagina e deve contenere indicazioni utili al browser per sapere quale linguaggio incontrerà nelle righe del codice. Per creare le pagine web può essere usato uno di questi tre linguaggi:

- 1.HTML 5.0
- 2.XHTML 1.0
- 3.HTML 4.01

Per ognuno di essi la dichiarazione è leggermente diversa, come nei seguenti esempi:

- 1.<!DOCTYPE html>
- 2.<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
- 3.<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">

Spieghiamo il significato di queste dichiarazioni:

- la parola chiave **!DOCTYPE** (notare il punto esclamativo)
- il nome dell'elemento root (radice) del documento (nelle pagine web il tag che contiene tutti gli altri tag è <html>)
- la parola chiave **PUBLIC**
- uno o due riferimenti al tipo di documento, cioè alle regole di sintassi. Il primo riferimento è il formal public identifier (FPI), il secondo (opzionale) è il formal system identifier (FSI), che è una URL.

Questa dichiarazione assomiglia ad un commento e quindi alcuni browser spesso lo ignorano, ma se il metadata "content-type" vale "text/html" non lo ignorano, anzi lo usano per individuare il numero della versione del linguaggio, come visto nei precedenti esempi.

Nel terzo esempio si dichiara di utilizzare il linguaggio HTML versione 4.01 (Strict), mentre nel secondo si dichiara di utilizzare la sintassi del linguaggio XHTML versione 1.0 (Strict).

Alcuni tag possono non essere usati nei documenti strettamente aderenti allo standard HTML (Strict) perché sono vecchi e deprecati, come <font>, ma si possono continuare ad usare in documenti in HTML (Transitional)

## HTML4, HTML5 e XHTML

Le differenze principali tra i linguaggi HTML4, HTML5 e XHTML sono:

- l'aggiunta di nuovi tag in HTML5.
  - che non si vedranno per motivi di tempo...
- il diverso modo di chiudere i tag in XHTML
  1. in XHTML i tag devono essere obbligatoriamente in minuscolo
  2. in XHTML i tag <p>, <a> (e altri) devono essere chiusi con </p> e </a>
  3. in XHTML i tag <meta>, <img>, <hr> (e altri) devono essere auto-chiusi:  
<meta />, <img /> e <hr />

## I metadati

I Metadati sono dati che descrivono altri dati.

Il tag <meta> contiene dati che descrivono il contenuto della pagina HTML e si usa all'interno del tag <head>.

Un metadato può specificare: l'autore della pagina, la descrizione del contenuto, le parole chiave usate dai motori di ricerca per catalogare le pagine web.

Un esempio in HTML 5:

```
<head>
  <meta name="description" content="Materiale per studiare" />
  <meta charset="UTF-8" />
</head>
```

Un esempio in XHTML:

```
<head>
  <meta name="description" content="Materiale per studiare" />
  <meta name="author" content="Mario Rossi" />
  <meta name="robots" content="index, follow" />
</head>
```

L'attributo **name** indica il nome del metadato e può valere: "author", "description", "keywords", "robots" e altri.

L'attributo **http-equiv** indica il nome dell' "header HTTP" e può valere: "content-type", "content-style-type", "expires", "refresh", "set-cookie".

In entrambi i casi, dopo gli attributi name o http-equiv, si usa l'attributo **content** che specifica il valore contenuto nel metadato.

## Codifica dei caratteri

Prima di iniziare a creare un file di testo, con qualsiasi estensione, ad esempio .htm, quando si salva si deve prima scegliere la codifica da adottare per trasformare i caratteri in numeri nella memoria del computer.

Ogni editor di testo "serio" ha questa funzione.

Il consorzio W3C (World Wide Web Consortium) raccomanda l'utilizzo di UTF-8 perché permette di essere usato per pagine in tutte le lingue.

Per un documento in HTML 5.0:

```
<head>
  <meta charset="UTF-8">
</head>
```

Nei documenti "text/html" il metadato "content-type" specifica anche il "charset" (codifica dei caratteri). Negli altri tipi di documento la codifica dei caratteri non dovrebbe essere specificata. Il metadato "content-type" potrebbe essere specificato in modi diversi a seconda del linguaggio usato:

Per un documento in HTML 4.01:

```
<head>
  <meta http-equiv="Content-Type"
        content="text/html; charset=utf-8" />
</head>
```

Oppure

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

Per un documento in XHTML 1.0:

```
<head>
  <meta http-equiv="Content-Type"
        content="text/html; charset=utf-8" />
</head>
```

Oppure

```
<head>
  <meta http-equiv="Content-Type"
        content="application/xhtml+xml" />
</head>
```

Il più importante degli header HTTP è il "content-type" (tipo di contenuto), simile al MIME Type usato nella posta elettronica.

Il MIME type (Multipurpose Internet Mail Extensions) è una specifica nata nelle comunicazioni di posta elettronica per formattare un contenuto di tipo non-ASCII (non testuale). Oggi il MIME Type è utilizzato in modo diffuso anche dai browser per poter ricevere sul web contenuti come audio, video o applicazioni. Nelle pagine web possono essere utilizzati, ad esempio, i seguenti tipi/sottotipi di contenuto (content-type):

*application/xhtml+xml*

*application/xml*

*text/html* (vedere <http://www.ietf.org/rfc/rfc2854.txt>)

*text/xml*

A questo punto è utile leggere di nuovo il primo esempio. Il tag <meta> specifica il metadato "Content-Type" e gli assegna uno dei valori che possono essere assegnati anche ad un "MIME type", come text/html.

## La lingua

**lang** e **xml:lang**, sono attributi standard che specificano il linguaggio usato nel testo contenuto in un elemento (tag) come <html>, <h1>, <p>, ecc.

Anche il tag <html> può usare questi attributi per specificare la lingua:

<html lang="en">

<html lang="en" xml:lang="en" xmlns="http://www.w3.org/1999/xhtml">

**xmlns** è un attributo che specifica il "namespace" di un documento xhtml: non è usato in HTML ma è obbligatorio in XHTML e deve avere il seguente valore:

<html xmlns="http://www.w3.org/1999/xhtml">

## Caratteri e simboli (detti Entities)

Ripetiamo velocemente cosa è necessario fare per creare una pagina web che non presenti problemi di codifica dei caratteri non standard:

- scegliere la codifica nell'editor prima di salvare il file
- indicare il linguaggio utilizzato (HTML o XHTML) e il numero della sua versione (in <!DOCTYPE>);
- indicare il tipo di contenuto e l'eventuale codifica (attributo in <meta>);
- indicare eventualmente la lingua utilizzata (attributo in <html>);

Inoltre bisogna porre attenzione alla rappresentazione di alcuni caratteri speciali come, ad esempio:

simbolo	descrizione (inglese)	prima rappresentazione	seconda rappresentazione
€	euro	&euro;	&#8364;
<	less than	&lt;	&#60;
>	greater than	&gt;	&#62;
à	small a, grave accent	&agrave;	
ì	small i, grave accent	&igrave;	

simbolo	descrizione (inglese)	prima rappresentazione	seconda rappresentazione
è	small e, grave accent	&egrave;	&#232;
é	small e, acute accent	&eacute;	
È	upper e, grave accent	&Egrave;	

## Download

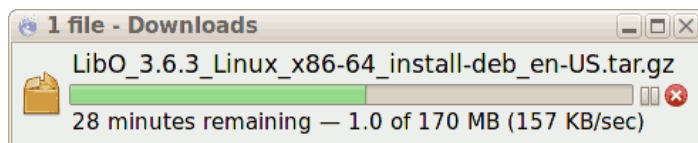
Chi naviga sul web ha letto molto spesso la parola *download*, che significa "scaricare".

Significa che si può ottenere una copia di qualcosa, come un documento, un file audio, o un programma, sul proprio computer.

Il download è un'operazione molto nota per chi naviga sul web, ma è allo stesso tempo un'operazione "delicata", nel senso che con il download si introduce qualcosa di "estraneo" sul proprio computer, senza conoscere bene chi ce lo fornisce. Questo potrebbe essere un rischio per la sicurezza software.

Di solito il download è un'operazione volontaria, cioè l'utente deve cliccare su qualcosa per far iniziare il download e deve decidere in quale cartella salvare l'oggetto scaricato.

A seconda delle dimensioni (in Byte), e a seconda della velocità del collegamento ad Internet (in Byte per



secondo), sarà necessario un certo tempo per il Download. Il completamento del download è mostrato dal completamento di una barra di progressione, come nella precedente figura.

Il download può essere avviato con un click su un collegamento ipertestuale, usando il tasto destro del mouse, come nella seguente figura (salva con nome):

Anche quando stiamo guardando la foto di una notizia sul web, potremmo vederla comparire lentamente. Questo significa che anche la foto della pagina web viene scaricata sul nostro computer, ma in questo caso, senza che noi avessimo fatto nessun click.



## File temporanei del browser

Ogni browser possiede una propria cartella dei file temporanei di Internet, che si riempie con tutto quello che viene visualizzato dal browser mentre navighiamo: immagini, testi, video, ecc.

Questa cartella (detta anche cartella cache del browser) permette di visualizzare le pagine visitate anche quando si è disconnessi da Internet (modalità offline).

Questa cartella si può trovare in diversi posti, a seconda del browser che utilizziamo.

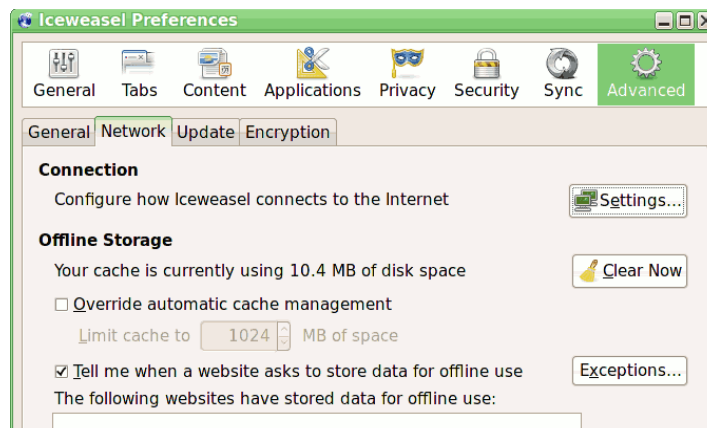
Attività: dove si trova questa cartella per il browser Microsoft Internet Explorer? Quella di Mozilla Firefox? Quanto spazio occupa nella memoria questa cartella? È possibile svuotarla? Attenzione: ci sono differenze anche per differenti versioni dello stesso browser.

Chi usa Mozilla Firefox può cercare la risposta a questa domanda, facendo una semplice ricerca su un motore di ricerca.

Firefox: **Menu Tools >**

Options > Advanced > Network > Offline Storage (Cache): "Clear Now"

Nella seguente figura è mostrata la finestra di dialogo che permette di configurare la gestione della cache dei file temporanei per Firefox.



## Cookies

La parola cookie significa letteralmente "biscotto", ma cosa c'entra un biscotto con il web?

Cosa sono i cookies?

I cookies sono in realtà dei semplici file di testo che contengono i dati dell'utente che naviga, per questo potrebbero anche essere utilizzati per violare la nostra privacy.

Dove si trovano i cookies?

I cookie si trovano nella cartella dei file temporanei di Internet, di cui si è già parlato.

Chi crea i cookies?

Il browser genera i cookies automaticamente (senza avvisarci) quando viene richiesto dal sito che stiamo visitando. Ogni browser può essere configurato per comportarsi in modo diverso, e avvisarci.

Per cosa sono utili?

Sono utili quando vogliamo essere riconosciuti automaticamente al nostro ingresso su un sito.

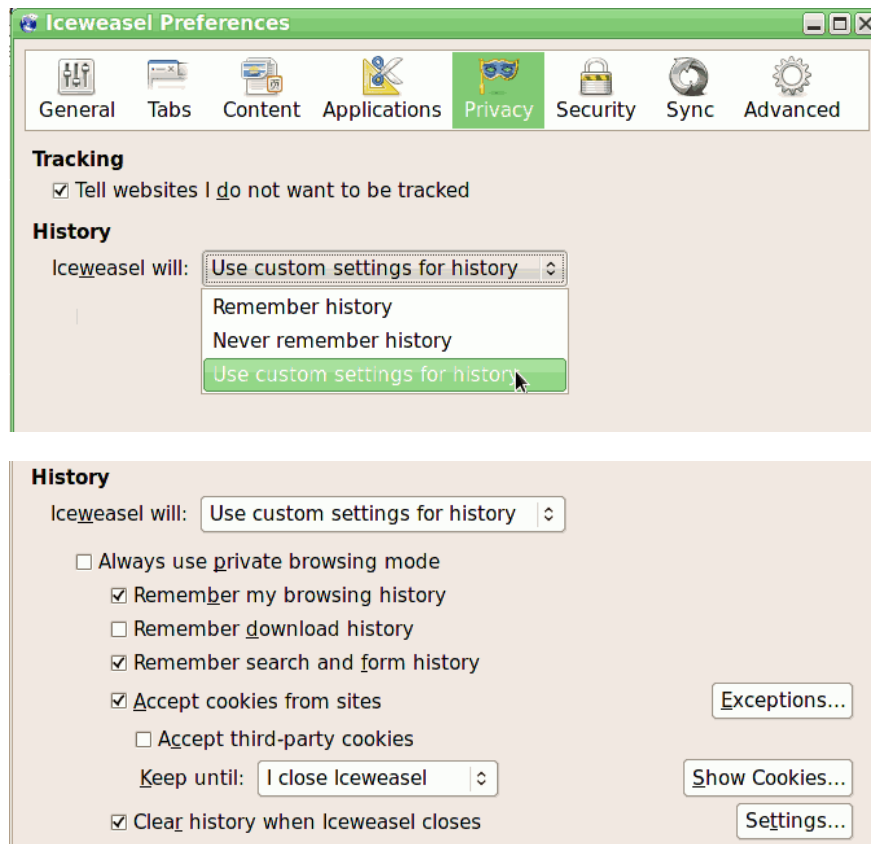
Usando i cookie si può evitare di digitare la password.

Per cosa sono dannosi?

Sono utilizzati anche da chi vende pubblicità su Internet per controllare il nostro comportamento e le nostre abitudini, quindi sono potenzialmente anche una minaccia alla nostra privacy.

Attività: è possibile eliminare i cookie, ma come si fa per Internet Explorer? E per Firefox?

Nelle seguenti figure si possono leggere le configurazioni possibili per tutelare la propria privacy con Firefox 10.0





## Inline frame <iframe>

In inglese "frame" vuol dire "cornice".

Questo tag consente di creare, nella pagina web, una cornice che contiene a sua volta un'altra pagina web.

```
<iframe src="./prima.htm">
</iframe>
```

Il tag può essere usato più volte e queste cornici possono essere usate per realizzare una pagina complessa composta a sua volta da diverse altre pagine.

```
<iframe src="./titolo.htm">
</iframe>
<iframe src="./contenuto.htm">
</iframe>
<iframe src="./piede_di_pagina.htm">
</iframe>
```

Per consentire all'utente di navigare aprendo anche i contenuti dei link all'interno di un <iframe> (inline frame) si devono usare gli attributi "target" e "name"

```
<a href="telefono.htm" target="cornice1"> telefono </a>
<iframe name="cornice1" src="./prima.htm">
</iframe>
```

L'attributo src ha la funzione di specificare la pagina iniziale da inserire dentro un iframe.

Alcuni vecchi browser non supportano i frame, quindi può essere utile inserire un avviso tra <iframe> e </iframe> che verrà visualizzato solo nei vecchi browser.

```
<iframe name="cornice1" src="./prima.htm">
    <p>avviso: aggiorna il tuo browser</p>
</iframe>
```

Altri possibili attributi sono width, height e scrolling

```
<iframe name="cornice1" src="./prima.htm" width="500" height="440"
scrolling="auto">
    <p>avviso: aggiorna il tuo browser</p>
</iframe>
```