

DA HTML A XHTML

Francesca Di Donato, [didonato\[at\]sp.unipi.it](mailto:didonato[at]sp.unipi.it)



L'evoluzione del Web: 1990-2007

Partito con una pagina ai primi anni 90, nel 2005 il WWW contava più di 100 milioni di siti, e 8 miliardi di pagine pubblicamente accessibili.

Perché?

Non ci sono limiti al numero di link e al numero di pagine che si possono aggiungere.

Per farlo, non bisogna coordinarsi con nessuna autorità centrale.

<<http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html>>

L'evoluzione del Web: 1990-2007

Partito con una pagina ai primi anni 90, nel 2005 il WWW contava più di 100 milioni di siti, e 8 miliardi di pagine pubblicamente accessibili.

Perché?

→ Non ci sono limiti al numero di link e al numero di pagine che si possono aggiungere.

Per farlo, non bisogna coordinarsi con nessuna autorità centrale.

<<http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html>>

L'evoluzione del Web: 1990-2007

Partito con una pagina ai primi anni 90, nel 2005 il WWW contava più di 100 milioni di siti, e 8 miliardi di pagine pubblicamente accessibili.

Perché?

- ➔ Non ci sono limiti al numero di link e al numero di pagine che si possono aggiungere.
- ➔ Per farlo, non bisogna coordinarsi con nessuna autorità centrale.

<<http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html>>

2007

Il web è fatto di

siti, pagine e **documenti**

noi (umani) lo vediamo
così

Login

Cerca

Titoli

Testo

Bollettino telematico di filosofia politica: **AnalisiLinguaggioPolitico2**

RecentChanges

FindPage

HelpContents

AnalisiLing...ioPolitico2

Modifica (modo testuale)

Modifica (modo grafico)

Informazioni

Allegati

Altre Azioni:

Analisi del linguaggio politico II

Facoltà di Scienze politiche dell'Università di Pisa, AA 2007-2008

Parte Seconda

Programma

Una lettura a scelta tra:

1. A. Turing, *Macchine calcolatrici e intelligenza*, in V. Somenzi (a cura di), *La filosofia degli automi*, Bollati Boringhieri, Torino 1965 (*Computing Machinery and Intelligence*, *Mind*, London N.S. vol. 59, 433-60, 1950, disponibile in rete al seguente url: <http://www.abelard.org/turpap/turpap.htm>).
2. V. Bush, "As we may think" (1945) *The Atlantic Monthly*, luglio 1945, vol. 176, No. 1, pp. 101-108. (Il testo in lingua originale: <http://www.theatlantic.com/doc/194507/bush>)
3. T. Berners-Lee, "The World Wide Web - Past, present, future. Exploring Universality" (2002): <http://www.w3.org/2002/04/Japan/Lecture.html>

Materiali didattici (in aggiornamento)

* Slide della prima lezione: [ALP201.pdf](#) [2,8 MB]

le macchine lo vedono così


```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
<html>
  <head>
    <meta http-equiv="Content-Type" content="text/html;charset=utf-8">
    <meta name="robots" content="noindex,nofollow">

    <title>AnalisiLinguaggioPolitico2 - Wiki del Bollettino telematico di filosofia politica</title>
    <script type="text/javascript" src="/wiki/common/js/common.js"></script>

    <script type="text/javascript">

      <!--// common functions
      // We keep here the state of the search box
      searchIsDisabled = false;

      function searchChange(e) {
        // Update search buttons status according to search box content.
        // Ignore empty or whitespace search term.
        var value = e.value.replace(/\s+/, "");
        if (value == "" || searchIsDisabled) {
          searchSetDisabled(true);
        } else {
          searchSetDisabled(false);
        }
      }

      function searchSetDisabled(flag) {
        // Enable or disable search
        document.getElementById('fullsearch').disabled = flag;
        document.getElementById('titlesearch').disabled = flag;
      }

      .....
      .....
      .....
```

i limiti di HTML

Un documento HTML è un file di testo che contiene: Testo + Markup

- Confonde l'aspetto di un documento e la semantica del testo
- Comprende la possibilità di overlapping


HTML

non permette

lo scambio di informazione strutturata
tra le macchine

XML

dialetto di SGML con HTML, è pensato per tenere distinti “data” e “documents”

- Struttura ad albero
- “well-formedness”: **i file devono essere ben formati**
- validità: **i file devono essere validi** 

NB. XML è **case sensitive!**

La struttura ad albero di XML

I file XML hanno una struttura ad albero, cioè

2. Non è possibile l'overlapping (sovrapposizione):

```
<html>  
<body>  
<p>testo <b>in  
grassetto</p>  
<p> altro testo e poi chiudo il  
grassetto</b>  
  
</body>  
</html>
```

1. Hanno necessariamente un elemento radice (root) che dev'essere unico (<html>)



File ben formati

Elementi

- descrivono i contenuti e la struttura del documento, es. p come paragrafo
- sono delimitati da tag di apertura e chiusura, “<” e “>”, es. <p>Questo è un paragrafo</p>
- il contenuto di un elemento può essere vuoto, es.

- tutto ciò che si trova all’interno di un commento non viene elaborato, es. <!-- Commento -->
- sono scritti tutti e solo in minuscolo (vale anche per gli attributi)

File ben formati

Attributi

- gli elementi possono contenere uno o più attributi, che sono termini aggiuntivi che definiscono con più precisione l'elemento, es. ` `
- il valore degli attributi è racchiuso tra virgolette (singole: 'uno', o doppie: "uno"). N.B. Le virgolette devono essere entrambe singole o entrambe doppie.
- Un attributo non può essere usato due volte all'interno dello stesso elemento
- Un elemento può contenere più di un attributo

File ben formati

Struttura: riepilogo

1. Annidamento

```
<html>  
  <head>  
    <title> Titolo </title>  
  </head>  
  <body>  
    <p>Paragrafo 1</p>  
    <p> Paragrafo 2</p>  
  </body>  
</html>
```

2. Relazioni gerarchiche e ordinali

Dati due elementi a e b, può darsi uno solo di questi casi:

1. a contiene b
2. a è contenuto da b
3. a precede b
4. a segue b

Esistono moltissimi dialetti di XML:

- XHTML,
- DocBook,
- Text Encoding Initiative (TEI)
-

**Noi vediamo
XHTML 1.0**

File validi

Un file XHTML è valido se conforme alla sua DTD (Document Type Definition)

La DTD definisce le regole di “dialetti” particolari

Esistono 3 DTD XHTML del W3C:

1. Strict: *Use this when you want really clean markup, free of presentational clutter. Use this together with Cascading Style Sheets.*

2. Transitional: *Use this when you need to take advantage of HTML's presentational features and when you want to support browsers that don't understand Cascading Style Sheets.*

3. Frameset: *Use this when you want to use HTML Frames to partition the browser window into two or more frames.*

File validi



```
xhtml1-strict.dtd
1 <!--
2 Extensible HTML version 1.0 Strict DTD
3
4 This is the same as HTML 4 Strict except for
5 changes due to the differences between XML and SGML.
6
7 Namespace = http://www.w3.org/1999/xhtml
8
9 For further information, see: http://www.w3.org/TR/xhtml1
10
11 Copyright (c) 1998-2002 W3C (MIT, INRIA, Keio),
12 All Rights Reserved.
13
14 This DTD module is identified by the PUBLIC and SYSTEM identifiers:
15
16 PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
17 SYSTEM "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd"
18
19 $Revision: 1.1 $
20 $Date: 2002/08/01 13:56:03 $
21
22 -->
23
24 <!------- Character mnemonic entities ----->
25
26 <!ENTITY % HTMLlat1 PUBLIC
27 "-//W3C//ENTITIES Latin 1 for XHTML//EN"
28 "xhtml-lat1.ent">
29 %HTMLlat1;
30
31 <!ENTITY % HTMLsymbol PUBLIC
32 "-//W3C//ENTITIES Symbols for XHTML//EN"
33 "xhtml-symbol.ent">
34 %HTMLsymbol;
35
36 <!ENTITY % HTMLspecial PUBLIC
37 "-//W3C//ENTITIES Special for XHTML//EN"
38 "xhtml-special.ent">
39 %HTMLspecial;
40
41
```

Noi useremo la
DTD
**XHTML 1.0-
Transitional**

NB

La DTD è
un file - si può
anche scaricare e
salvare in locale
(cfr. img a
sinistra)-

File validi

Un collegamento alla DTD dev'essere inserito in apertura del file:

```
<!DOCTYPE html  
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
```

```
<html>
```

```
</html>
```

Il resto del file
è un file html
normale

File validi



```
<!DOCTYPE html
PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//
EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-
transitional.dtd">
```

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

```
<head>
```

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

```
<title>Indice</title>
```

```
</head>
```

```
<body>
```

...



Bisogna
inserire anche
gli elementi
qui in
grassetto

XHTML: Elementi obbligatori

Questo è un modello “minimo” di documento XHTML:

```
<!DOCTYPE Doctype goes here>
<html xmlns="http://www.w3.org/1999/xhtml">

<head>
<title>Titolo</title>
</head>

<body> .....
</body>

</html>
```

- Un documento XHTML **deve avere la dichiarazione DOCTYPE.**

- Gli elementi **html**, **head** e **body** devono essere **presenti** (e **head** deve contenere l'elemento **title**).

NB: xmlns sta per xml namespace

Riepilogando

- gli elementi singoli vengono scritti diversamente da quanto avveniva in HTML: es. `
` diventa `
`
- Un documento XHTML **deve avere la dichiarazione DOCTYPE e, sotto, il namespace(*) di xml (xmlns).**
- Gli elementi **html**, **head** e **body** **devono** essere **presenti** (e **head** deve contenere l'elemento **title**).
- i tag devono essere scritti tutti in **minuscolo**.

(*) I **namespace** xml offrono un metodo semplice per qualificare i nomi di elementi e attributi usati nei documenti xml associandoli con spazi di denominazione identificati da **uri**.

XHTML

TAG	DESCRIZIONE	DTD
<!--...-->	commento	stf
<!DOCTYPE>	document type	stf
<a>	ancora	stf
<acronym>	acronimo	stf
<address>	indirizzo	stf
	grassetto	stf
 	a capo	stf
<button>	bottone	stf
<cite>	citazione	stf

<<http://www.w3schools.com/tags/default.asp>>

All'indirizzo: <<http://www.w3schools.com/tags/default.asp>>

trovate tutti gli elementi (X)HTML e i relativi attributi, con una descrizione precisa dell'uso e alcuni esempi. Utilizzate la pagina per costruire il vostro primo sito Web. Non dimenticate, alla fine, di validarlo!!

TAG	DESCRIZIONE	DTD
		stf
		stf
		stf
		stf

Un esempio di sito in XHTML...

il codice può essere copiato!!

La teoria femminista: una bibliografia

http://bfp.sp.unipi.it/biblio added/home.htm

università di Trento

www.sp.unipi.it - Ho... Isis is working Deutsches Wörterbuc... Unipi-Eprints - Arch... Lo Spettacolo: guida t... Françoise Rétif: Mod...

Le... H... Se... CS... D... Pri... Ca... Pri... D... Ch... Po... La...

Bollettino telematico di filosofia politica

La teoria femminista

Una bibliografia

Percorsi di lettura
Ipertesti di ricerca **New!**
Testi on-line **New!**
Indice alfabetico dei titoli
Mappe

Ipertesto a cura di Francesca Di Donato (france[at]sss sup.it) - Ultimo aggiornamento 2 aprile 2004

Valid XHTML 1.0!

Completato

<http://bfp.sp.unipi.it/biblio added/>

Il codice XHTML dello stesso sito... (dal browser: visualizza -> sorgente pagina)

```
view-source: - Sorgente di: http://bfp.sp.unipi.it/biblio added/home.htm

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en"><head><title>La teoria femminista: una bibliografia</title>

  <meta http-equiv="Content-Type" content="text/html; charset=ISO-8859-1">
  <meta name="author" content="francesca di donato">
  <meta name="Description" content="Bibliografia di teoria femminista: testi e ipertesti">
  <meta name="Keywords" content="femminismo, differenza, genere, filosofia politica, feminism">
  <link rel="stylesheet" type="text/css" href="bfp_biblio added.css"></head>

  <body bgcolor="#000066">

  <table bgcolor="#000066" border="0" cellpadding="2" cellspacing="0" width="100%">
    <tbody><tr>
      <td align="center">

  
  </td>
  <td align="center">
  

  <br>

  <a href="http://purl.org/hj/bfp/"><font color="#ff9900" size="-2"><strong>bfp</strong></font></a>
  </td>
  </tr>
  </tbody></table>
  <table bgcolor="#000066" border="0" cellpadding="2" cellspacing="0" width="95%">
  <tbody><tr>
  <td>
  <br>
  <h1>La teoria femminista
  <br><font size="-1">Una bibliografia</font>
  </h1>
  <br>
  <br>
  <p>
```

<http://bfp.sp.unipi.it/biblio added/>

esercizio 9

Trasformate la vostra pagina index.html in un file xhtml transitional. Nel farlo, controllate che l'albero html sia ben formato (no overlapping!).

Verificare se un file è valido:

<http://validator.w3.org>

The screenshot shows a browser window titled "The W3C Markup Validation Service" with the URL "http://validator.w3.org/". The browser's address bar shows "w3c xhtml". The page features a navigation menu with links for Home, About..., News, Docs, Help & FAQ, Feedback, and Donate. The main content area includes a description of the service, a "Validate Your Markup" section with three options: "Validate by URL", "Validate by File Upload", and "Validate by Direct Input".

W3C QUALITY Assurance **Markup Validation Service v0.7.4**

Home About... News Docs Help & FAQ Feedback Donate

This is the W3C Markup Validation Service, a free service that checks Web documents in formats like HTML and XHTML for conformance to W3C Recommendations and other standards.

if you wish to validate specific content such as RSS/Atom feeds or CSS stylesheets or to find broken links, there are other validators and tools available.

Validate Your Markup

Validate by URL

Address:

Enter the [URL](#) of the page you want to check. Advanced options are available from the [Extended Interface](#).

Validate by File Upload

Local File:

Select the file you want to upload and check. Advanced options are available from the [Extended File Upload Interface](#).

Note: file upload may not work with Internet Explorer on some versions of Windows XP Service Pack 2, see our [information page](#) on the W3C QA Website.

Validate by Direct Input

Input the markup you would like to validate in the text area below:

Trova: Trova successivo Trova precedente Evidenzia M Maiuscole/minuscole

esercizio 10

Validate index.html al validatore w3c:

<http://validator.w3.org>

NB. E' molto importante imparare a leggere i messaggi di errore del validatore e a correggere il file di conseguenza.

NB2. Il validatore ha sempre ragione... sempre sempre!

esercizio 11

Trasformate anche due.html e tre.html in documenti xhtml transitional.

esercizio 12

Validate questo file:

[http://bfp.sp.unipi.it/moinwiki/LaboratorioScrittura/
esercizio12.html](http://bfp.sp.unipi.it/moinwiki/LaboratorioScrittura/esercizio12.html)

e correggete gli errori (per verificare di esserci riusciti, passatelo di nuovo al validatore)

esercizio 13

Codificate in XHTML il file:

[http://bfp.sp.unipi.it/moinwiki/LaboratorioScrittura/
esercizio13.rtf](http://bfp.sp.unipi.it/moinwiki/LaboratorioScrittura/esercizio13.rtf)

poi validatelo al validatore e correggete gli errori (per verificare di esserci riusciti, passatelo di nuovo al validatore)

Biblio-linkografia

XHTML

La raccomandazione del W3C

<http://www.w3.org/TR/xhtml1/>

Tutorial XHTML (sul sito del W3C)

<http://www.w3schools.com/xhtml/>

Validatore XHTML e CSS del W3C

<http://validator.w3.org>

HTML Tidy

<http://www.w3.org/People/Raggett/tidy/>

Tutti gli elementi (X)HTML (sul sito del W3C)

<http://www.w3schools.com/tags/default.asp>

Arachnophilia (Editor HTML) - **consigliato**

<http://www.arachnoid.com/arachnophilia/>

Amaya (Editor HTML multiplatforma)

<http://www.w3.org/Amaya/Overview.html>