

CHE COS'È L'XML?

- ❑ XML: eXtensible Markup Language è un **linguaggio** estensibile realizzato per poter utilizzare in modo semplice i documenti strutturati su supporto digitale.
- ❑ è uno strumento potente e versatile per la creazione, memorizzazione e distribuzione di documenti digitali
- ❑ Il linguaggio è definito “estensibile” in quanto, più che essere un linguaggio, è un vero e proprio metalinguaggio, utilizzabile per creare linguaggi di marcatura specifici.

LE ORIGINI

- ❑ La prima bozza di XML risale al novembre 1996 quando fu formato un gruppo di lavoro con l'incarico di definire un linguaggio a markup estensibile di uso generale.
- ❑ XML deriva da **SGML** e cioè un metalinguaggio definito come standard ISO (ISO 8879:1986 SGML) nato con lo scopo di definire linguaggi da utilizzare per la stesura di testi destinati ad essere trasmessi e archiviati con strumenti informatici, ossia per la stesura di documenti in forma leggibile da computer (machine readable form)
- ❑ XML nasce come un sottoinsieme semplificato di SGML orientato all'utilizzo su World Wide Web.

SGML

- ❑ **SGML (Standard Generalized Markup Language)** è il padre dell'XML ma anche dell'HTML.
- ❑ E' un linguaggio potente e flessibile ma dotato di una struttura troppo pesante per applicazioni WEB.
- ❑ Sono obbligatori un DTD, uno StyleSheet e la validazione del documento.

COME SI PRESENTA UN DOCUMENTO XML

- ❑ Uno degli obiettivi di progettazione di XML è che sia in un formato leggibile dall'uomo, non deve quindi essere in un formato binario.
- ❑ XML è un **linguaggio a marcatori** (markup) costituito da istruzioni, definite **tag** o **marcatori**, che descrivono la struttura e la forma di un documento.
- ❑ I marcatori vengono inseriti all'interno del documento ed ognuno di essi identifica un elemento o componente del documento.
- ❑ XML utilizza la codifica dei caratteri Unicode e sia il testo che i marcatori sono memorizzati in formato ASCII.

ESEMPIO DI XML

❓ `<documento>` ← **MARCATORE DI INIZIO**

`<corpo>`
Testo del primo documento
`</corpo >`

`</documento>` ← **MARCATORE DI FINE**

Ogni tag di apertura deve avere un corrispondente tag di chiusura.

Le maiuscole e le minuscole sono interpretate differentemente, pertanto il tag `<nome>` è diverso da `<Nome>` e da `<NOME>`; i tag HTML si scrivono in maiuscolo e quelli XML in minuscolo.

CARATTERISTICHE DELL'XML

- ❑ XML è indipendente dal tipo di piattaforma hardware e software su cui viene utilizzato.
- ❑ può essere archiviato su qualsiasi tipo di supporto digitale
- ❑ può essere visualizzato su qualsiasi dispositivo output
- ❑ Può essere facilmente trasmesso via Internet tramite i protocolli HTTP, SMTP, FTP.
- ❑ Inoltre, sono disponibili varie applicazioni e librerie open source per la manipolazione di dati in formato XML basate su diversi linguaggi di programmazione (Java, PHP, Perl, Python...)
- ❑ Un'applicazione in grado di elaborare dati in formato XML è detta **elaboratore XML**

LINGUAGGIE GRAMMATICHE

- ❑ Per definire un linguaggio abbiamo bisogno di definire una **grammatica**: un insieme di regole che indica quali elementi possono essere utilizzati e con che struttura è possibile comporre documenti.
- ❑ Se un documento XML rispetta le regole definite da una grammatica, è detto **valido**.

DOCUMENTI BEN FORMATI E DOCUMENTI VALIDI

- ❑ In XML ci sono regole **sintattiche** (o meglio meta-sintattiche): come dobbiamo scrivere le informazioni all'interno dei documenti
- ❑ Ci possono essere anche regole **semantiche**: cosa possiamo scrivere in un documento XML
- ❑ Un documento XML che rispetta le regole sintattiche viene detto ben formato (**well-formed**)
- ❑ Un documento XML che rispetta le regole sintattiche e le regole semantiche si dice valido.
- ❑ **Un documento ben formato può non essere valido mentre un documento valido è necessariamente ben formato.**

Affinchè un documento XML sia ben formato

- ❑ deve contenere una dichiarazione corretta;
- ❑ Il corpo deve avere un unico elemento radice;
- ❑ Ogni elemento deve avere un tag di apertura e uno di chiusura;
- ❑ In omide i tag di apertura e chiusura devono coincidere;
- ❑ I valori attribuiti devono essere sempre racchiusi tra singoli o doppi apici.

STRUTTURA DI UN DOCUMENTO XML

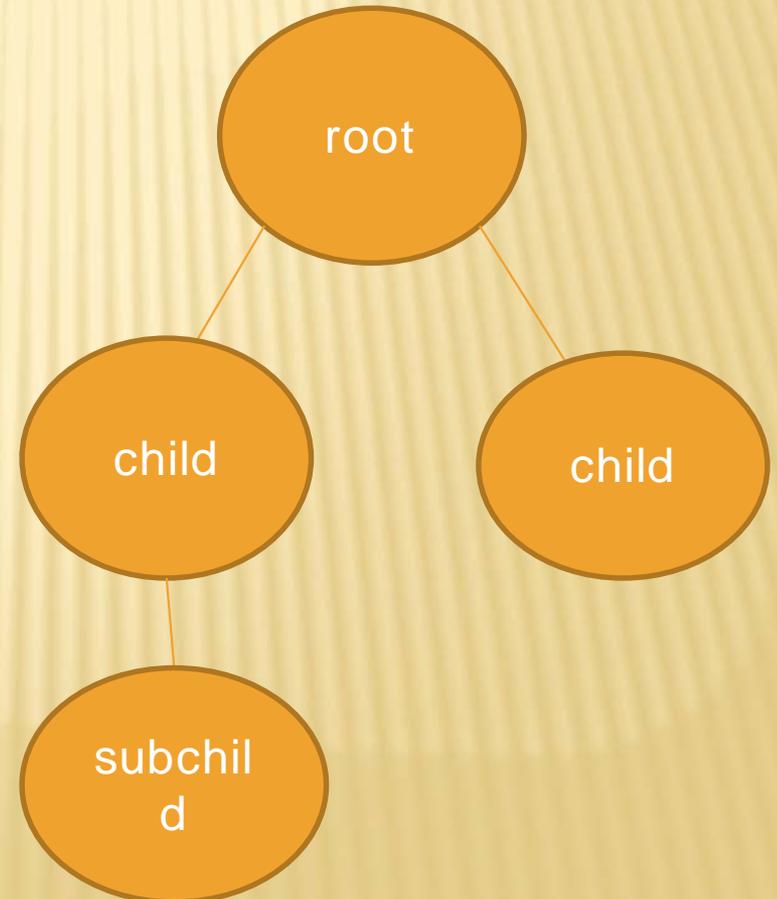
- ❑ Un documento XML è strutturato in modo gerarchico ed è composto da vari **elementi**.
- ❑ Un elemento può contenere un frammento di testo oppure altri elementi (**sotto-elementi**)
- ❑ Ad un elemento posso essere associate anche informazioni descrittive chiamate **attributi**.
- ❑ Inoltre, gli elementi sono organizzati ad un albero con radice **root**.
- ❑ Ogni documento XML può essere rappresentato come un albero : **document-tree**.

STRUTTURA LOGICA E FISICA

Struttura logica

```
<root>  
  <child>  
    <subchild>  
    ...  
  </subchild>  
</child>  
<child>  
...  
</child>  
</root>
```

Struttura fisica



PROLOGO: XML DECLARATION

- ❑ Ogni documento XML inizia con un prologo che contiene una dichiarazione di versione (XML Declaration)
- ❑ Forme di XML declaration:
`<?xml version="1.0"?>`
`<?xml version="1.0" encoding="UTF-8"?>`
- ❑ Contiene informazioni sulla versione (per ora solo 1.0) e sul set di caratteri: ASCII (set di caratteri a 7 bit);
- ❑ UTF-8 (unicode a 8 bit);
- ❑ ISO-8859-1 (Set Latin 1 con lettere accentate)

COMMENTI

Dopo il prologo vi è il corpo (l'XML vero e proprio).

Sia nel prologo che nel corpo possono apparire dei commenti che sono utili per spiegare la struttura del documento XML e commentare parti del documento durante le fasi di sviluppo e di test del software.

I commenti, però, non vengono mostrati dal browser ma sono visibili solo a chi guarda il codice sorgente del documento XML.

`<!-- Questo è un commento -->`

ELEMENTIE TAG

- ❓ Un elemento è un testo racchiuso tra uno start tag e un end tag.
- ❓ Uno **start tag** è costituito da un nome più eventuali attributi racchiusi in simboli “<“, “>”
- ❓ Un **end tag** è costituito dallo stesso nome dello start tag racchiuso da “</”, “>”
- ❓ Un **tag vuoto** è rappresentabile come:
❓ <TagName attribute-list />
- ❓ **I tag non possono avere nomi che iniziano per XML, XMI, Xml, xml...**

ATTRIBUTI

- ❑ Ad ogni elemento possono essere associati uno o più **attributi** che ne specificano ulteriori caratteristiche
- ❑ Gli attributi XML sono caratterizzati da:
 - ❑ un **nome**
 - ❑ da un **valore**

ESEMPIO ELEMENTIO ATTRIBUTI

```
<libro isbn="13 24 AX" titolo="L'amica geniale" />
```

```
<libro isbn="13 24 AX">  
<titolo>L'amica geniale</titolo>  
</libro>
```

```
<libro>  
<isbn>13 24 AX</isbn>  
<titolo>L'amica geniale</titolo>  
</libro>
```

XML NELL'AMBITO DELLE BIBLIOTECHE

```
<?xml version="1.0" ?>
```

```
<biblioteca>
```

```
  <catalogo>
```

```
    <sezione nome="Fantasy">
```

```
      <libro>
```

```
        <titolo>il signore degli anelli vol.1 </titolo>
```

```
        <autore>J.J.R Tolkien </autore>
```

```
        <isbn>458-89-203-4564-2 </isbn>
```

```
        <casa editrice>Bompiani </casa editrice>
```

```
        <disponibile>si </disponibile>
```

```
      </libro>
```

```
    </sezione>
```

```
    <sezione nome="Classici">
```

```
      <libro>
```

```
        <titolo>La Divina Commedia </titolo>
```

```
        <autore>Dante Alighieri </autore>
```

```
        <isbn>978-88-201-3789-2 </isbn>
```

```
        <casa editrice>Loesher edizioni </casa editrice>
```

```
        <disponibile>si </disponibile>
```

```
      </libro>
```

```
    </libro>
```

```
      <titolo>IMalavoglia </titolo>
```

```
      <autore>Verga </autore>
```

```
      <isbn>678-201-3789-2 </isbn>
```

XML NELL'AMBITO DELLE BIBLIOTECHE

```

        <casa editrice> Felt rinelli edizioni </casa editrice>
        <disponibile> si </disponibile>
    </libro>
</sezione>
</catalogo>
<personale>
    <lavoratore>
        <nome> Alessandro Rossi </nome>
        <data di nascita> 23-09-1980 </data di nascita>
        <email> ales.rossi123@gmail.com </email>
        <collocazione> box informazioni </collocazione>
        <ordinario> si </ordinario>
    </lavoratore>
    <lavoratore>
        <nome> Carla Bianchi </nome>
        <data di nascita> 05-07-1994 </data di nascita>
        <email> c.bianchi2@gmail.com </email>
        <collocazione> archivio </collocazione>
        <ordinario> no </ordinario>
    </lavoratore>
</personale>
</biblioteca>
```