



XML: extensible markup language

È uno strumento che consente la creazione, la memorizzazione e la distribuzione di documenti digitali

Usa la codifica dei caratteri Unicode

È un linguaggio a marcatori (markup) composto da istruzioni, definiti tag o marcatori, i quali descrivono la struttura e la forma di un documento

È case sensitive: distingue fra maiuscole e minuscole

Sia il testo, sia i marcatori sono memorizzati in formato ASCII: codice per la codifica di caratteri

CARATTERISTICHE



È indipendente dal tipo di piattaforma hardware e software su cui viene utilizzato



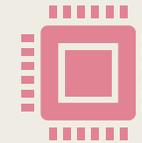
È indipendente dai dispositivi di archiviazione e visualizzazione



È uno standard di pubblico dominio

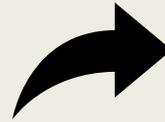


ogni software “conforme a XML” è in grado di gestire dati in formato XML



Un'applicazione in grado di elaborare dati in formato XML viene definita elaboratore XML

Metalinguaggi e grammatiche



XML è un metalinguaggio:
definisce un insieme regole (meta)sintattiche attraverso le
quali è possibile descrivere formalmente un linguaggio di
markup, detto applicazione XML

Ogni applicazione XML:
eredita un insieme di caratteristiche sintattiche comuni
definisce una sua sintassi formale
è dotata di una semantica

Regole



SINTATTICHE: COME SCRIVERE LE
INFORMAZIONI ALL'INTERNO DEL
DOCUMENTO



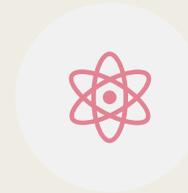
SEMANTICHE: COSA SCRIVERE IN
UN DOCUMENTO

- Un documento XML che rispetta le regole sintattiche si dice ben formato (well-formed)
- Un documento XML che rispetta le regole sintattiche e le regole semantiche si dice valido

DOCUMENTO XML



È strutturato in modo gerarchico



È composto da elementi



- Rappresenta una componente logico del documento
- Può contenere sottoelementi
- Possono essere associate informazioni descrittive (attributi)
- Gli elementi possono essere organizzati ad albero (document tree)

STRUTTURA FORMALE



Prologo: contiene dichiarazione XML ed il riferimento (opzionale) ad altri documenti.



Il prologo inizia con XML declaration: informazioni sulla versione e sui caratteri



Corpo: documento XML



Commenti: possono apparire ovunque in un documento; utili per spiegare e commentare parti del documento

Tag

Ogni elemento è racchiuso in uno start tag e un end tag

Start tag: costituito da un nome, eventuali attributi. Racchiusi dai simboli "<" e ">"

End tag: costituito da un nome

Attributi

Ad ogni elemento posso essere associati uno o più attributi

Sono caratterizzati da un nome

Sono caratterizzati da un valore

DTD: document type definition

Costituisce uno strumento di validazione per i documenti XML

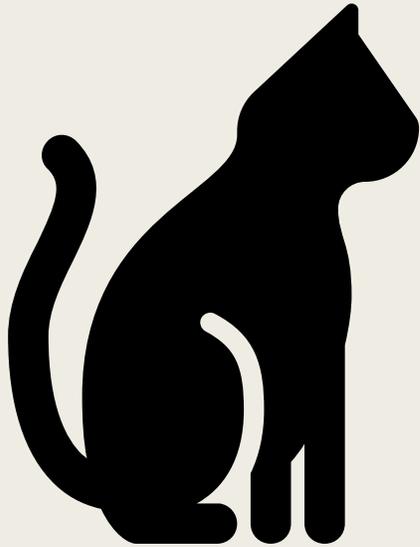
È costituito da un elenco di dichiarazioni che costituiscono il documento

Le dichiarazioni definiscono: gli elementi strutturali, il modello di contenuto e la lista degli attributi di ogni elemento

Il contenuto può essere di quattro tipi: empty, any, children mixed

Ogni elemento può essere indicato solo una volta

ESEMPIO DTD: CATALOGAZIONE GATTI



```
<!element catalogo (gatti*)>  
<!element gatto  
    (razza+, nome gatto?, nome padrone?, anni+, pelo?)  
    <!element razza (#PCDATA)>  
    <!element nome gatto (#PCDATA)>  
    <!element nome padrone (#PCDATA)>  
    <!element anni (#PCDATA)>  
    <!element pelo (#PCDATA)>
```

ESEMPIO XML: CATALOGAZIONE GATTI



```
<? xml version= "1.0"?>
<!DOCTYPE catalogo system "catalogo.dtd">
<catalogo>
  <gatto>
    <razza> Bombay </razza>
    <nome gatto> Lov </nome gatto>
    <nome padrone> Luca </nome padrone>
    <anni> 3 </anni>
    <pelo> pelo corto </pelo>
  </gatto>
  <gatto>
    <razza> Devon Rex </razza>
    <nome gatto> Ria </nome gatto>
    <nome padrone> Marco </nome padrone>
    <anni> 6 </anni>
    <pelo> pelo corto </pelo>
  </gatto>
  <gatto>
    <razza> Don Sphynx </razza>
    <nome gatto> Gi </nome gatto>
    <nome padrone> Laura </nome padrone>
    <anni> 8 </anni>
    <pelo> senza </pelo>
  </gatto>
  <gatto>
    <razza> Ragdoll </razza>
    <nome gatto> El </nome gatto>
    <nome padrone> Maria
    <anni> 5 </anni>
    <pelo> semilungo>
  </gatto>
  <gatto>
    <razza> Persiano </razza>
    <nome gatto> Pers </razza>
    <nome padrone> Anna </nome padrone>
    <anni> 5 </anni>
    <pelo> lungo </pelo>
  </gatto>
</catalogo>
```

Esiste una corrispondenza diretta fra struttura fisica e struttura logica (tree)

