



Università degli Studi di Ferrara  
Facoltà di Scienze MM FF NN  
Corso di Laurea in «*Scienze e Tecnologie per i Beni Culturali*»

AA 2010-2011

# INFORMATICA

Prof. Giorgio Poletti  
[giorgio.poletti@unife.it](mailto:giorgio.poletti@unife.it)

# XML: perché? XML: ortografia e sintassi

«Non hai veramente capito qualcosa  
fino a quando non sei in grado di  
spiegarlo a tua nonna.»

*(Albert Einstein)*

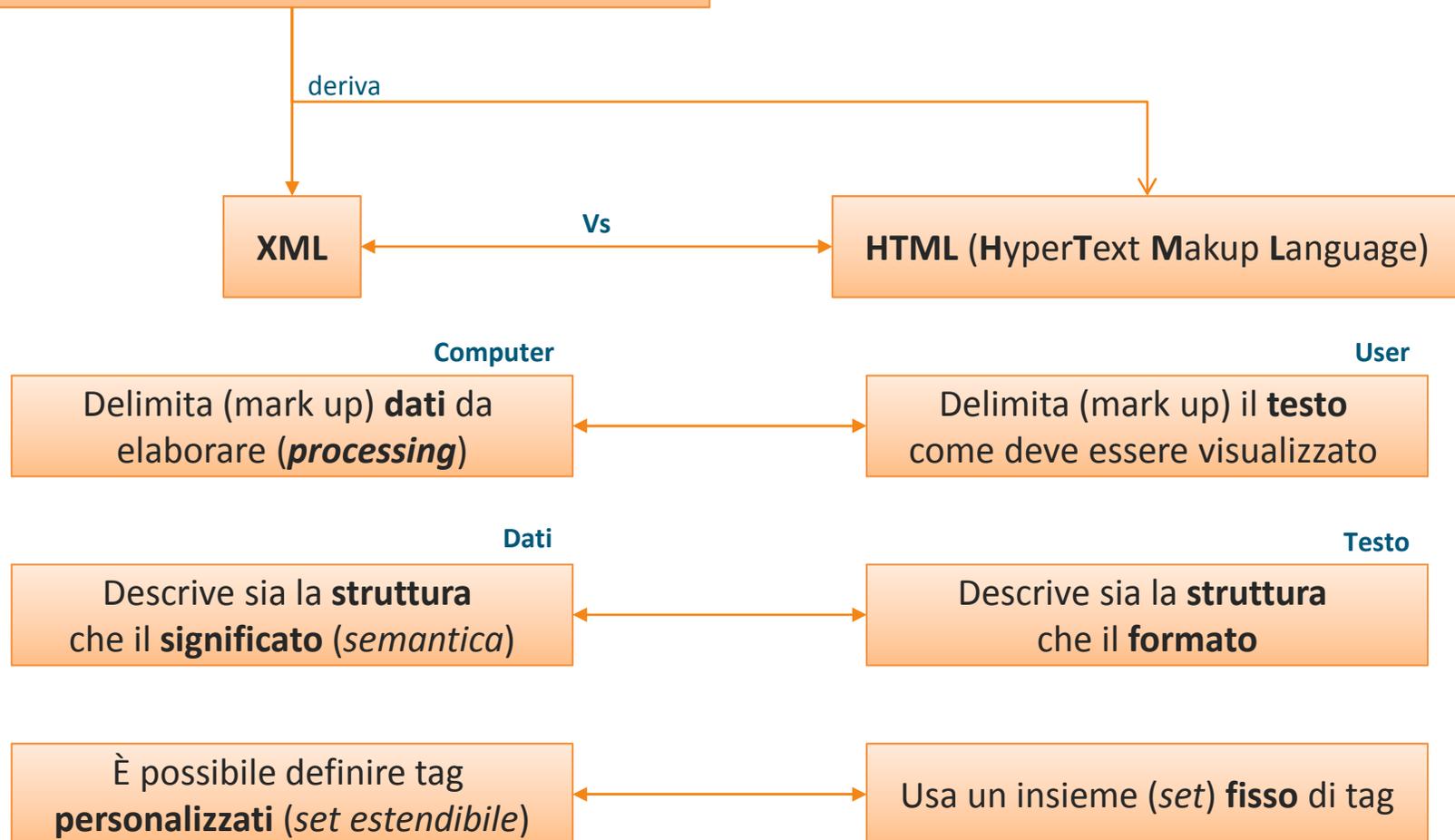


# XML: perché?

## XML: ortografia e sintassi

### XML – NOTE GENERALI

SGML (Standard Generalized Markup Language)

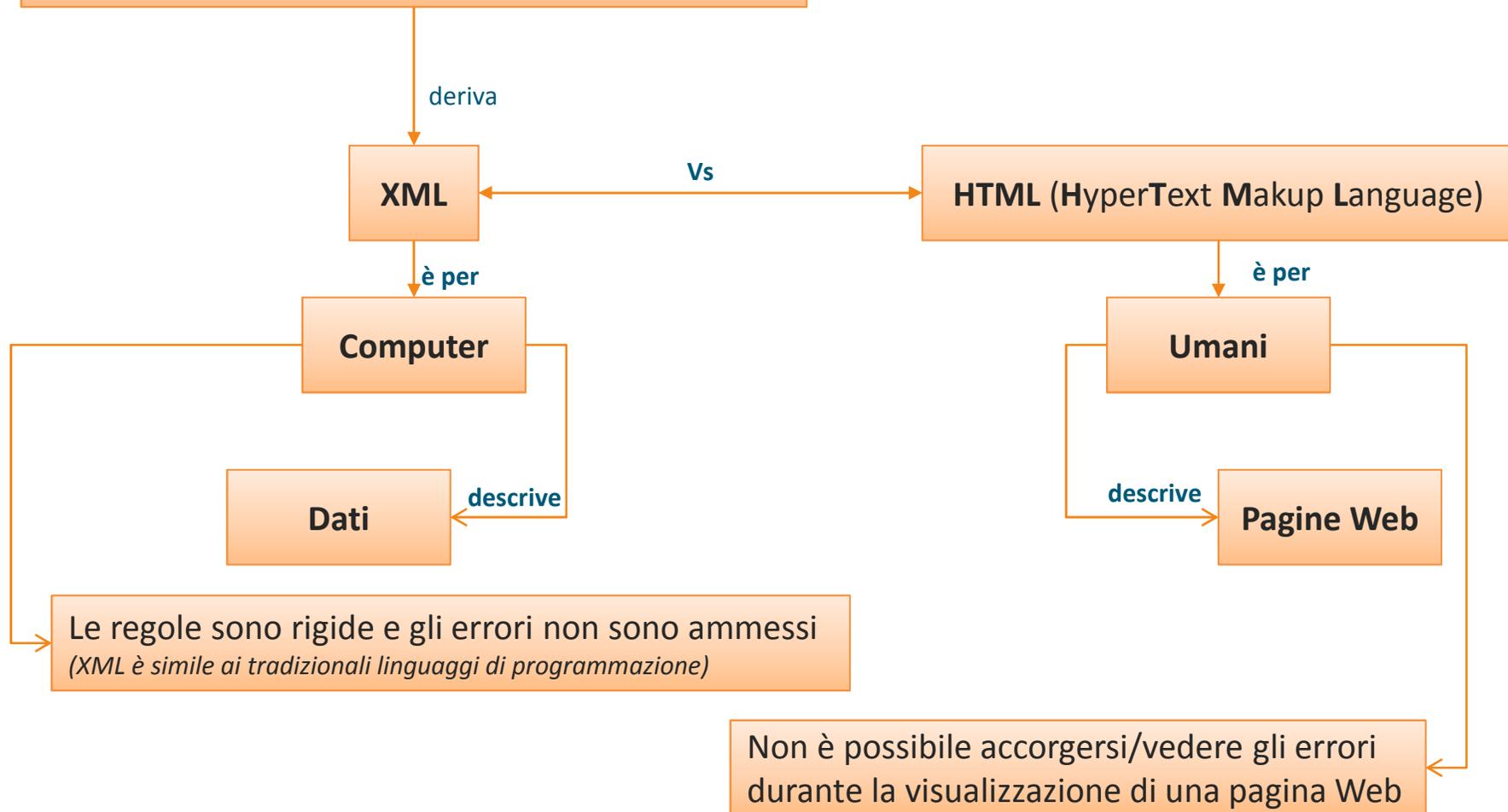


# XML: perché?

## XML: ortografia e sintassi

### XML – NOTE GENERALI

SGML (Standard Generalized Markup Language)



# XML: perché?

## XML: ortografia e sintassi

### XML – TECNOLOGIE

#### **DTD (Document Type Definition) e XML Schemas**

utilizzati per definire i **tag** utilizzabili nel documento XML, definisce i tag e i loro attributi per l'applicazione

#### **XSLT (eXtensible Stylesheet Language Transformations) e XPath**

utilizzati per trasformare tra diversi formati i dati di un XML

#### **CSS (Cascading Style Sheets)**

utilizzato per descrivere diversi formati da utilizzare in una pagina HTML o XML

#### **DOM (Document Object Model), SAX (Simple API<sup>1</sup> for XML), e JAXP (Java API for XML Processing)**

sono le all APIs per il parsing (controllo sintattico/grammaticale) di file XML

<sup>1</sup>API (*Application Programming Interface - Interfaccia di Programmazione di un'Applicazione*)

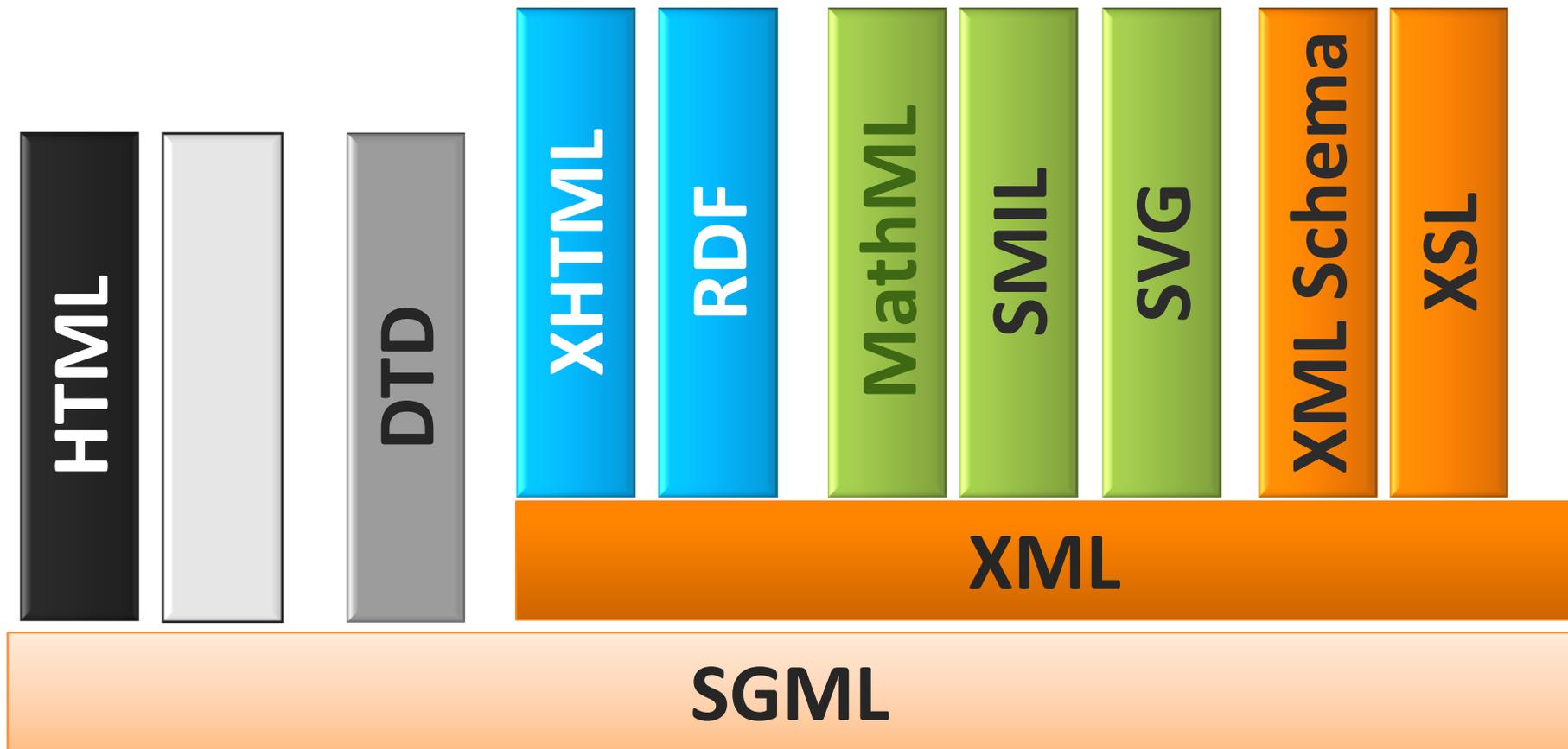
# XML: perché?

## XML: ortografia e sintassi

### XML – TECNOLOGIE

Java = portabilità dei programmi

XML = portabilità dei dati



# XML: perché?

# XML: ortografia e sintassi

## XML – ALFABETO E TAG

### FORMATO DI UN TAG (CASE SENSITIVE)

- **APERTURA**

<tagname attributo1='valore' ... attributo1='valore' >

NOME E 'VALORE' ATTRIBUTO

- **CHIUSURA**

</tagname>

NOME DEL TAG (non possono iniziare con un numero, caratteri speciali e contenere spazi)

### ALFABETO

a - z

0 - 9

A - Z

CARATTERI SPECIALI: / \ | : , . ; & ? ! < > = " "

# XML: perché?

# XML: ortografia e sintassi

## XML – STRUTURA DEL DOCUMENTO

### DOCUMENTO XML

## PROLOGO

DEFINIZIONE DI:

- VERSIONE DI XML
- DEFINIZIONE DELLE ENTITÀ (*entity*)
- DEFINIZIONE DEL TIPO DI DOCUMENTO (DTD o XML SCHEMA)

## COMPONENTI DEL DOCUMENTO

- TAG E ATTRIBUTI
- CDATA (**C**haracter **D**ATA)
- ENTITÀ (*entity*)
- **PI** (**P**rocessing **I**nstructions)
- COMMENTI

# XML: perché?

## XML: ortografia e sintassi

### XML – PROLOGO

```
<?xml version='1.0' encoding='ISO-8859-1' standalone='yes'?>
```

? Identifica le **PI (Processing Instruction)** o *direttive*

**version** attributo **OBBLIGATORIO** (*versione di XML*)

**encoding** attributo **OPZIONALE** (set di caratteri, default UTF-8)

**standalone** attributo **OPZIONALE** (definisce se viene fatto riferimento a un file esterno contenente la definizione di entità e struttura del documento, default yes)

*Il prologo può contenere definizione del documento e di entità*

# XML: perché?

## XML: ortografia e sintassi

### XML – PROLOGO

*Dichiarazione della struttura del documento*

**<!DOCTYPE root [DTD]>**

**! identifica i comandi DTD (scritti sempre in MAIUSCOLO)**

*Forma possibili per diversi [DTD]*

**<!DOCTYPE root SYSTEM URL>**

**Indirizzo privato**

**<!DOCTYPE root SYSTEM 'dir/file.dtd'>**

**<!DOCTYPE root SYSTEM 'http://indi.it/file.dtd'>**

# XML: perché?

## XML: ortografia e sintassi

### XML – PROLOGO

*Dichiarazione della struttura del documento*

**<!DOCTYPE root [DTD]>**

**! identifica i comandi DTD (scritti sempre in MAIUSCOLO)**

*Forma possibili per diversi [DTD]*

**<!DOCTYPE root PUBLIC FPI (Formal Public Identifier)- URL>**

***Indirizzo pubblico***

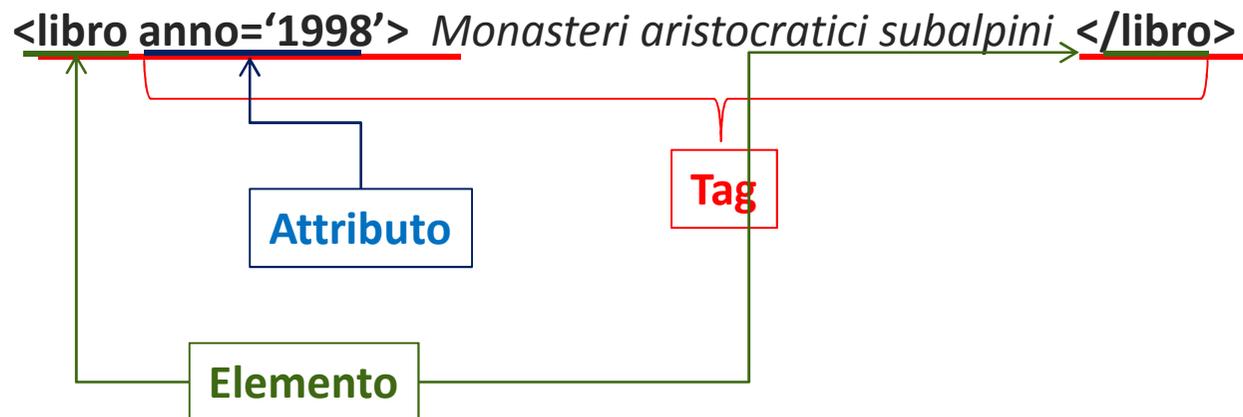
**<!DOCTYPE root PUBLIC '-//W3C//Transitional//EN' 'http://www.w3.org/TR?xhtml/file.dtd**

# XML: perché?

## XML: ortografia e sintassi

### XML – COMPONENTI DEL DOCUMENTO

- **TAG E ATTRIBUTI**
- CDATA (Character DATA)
- ENTITÀ (*entity*)
- PI (*Processing Instructions*)
- COMMENTI



<![CDATA[

# XML: perché?

## XML: ortografia e sintassi

### XML – COMPONENTI DEL DOCUMENTO

- TAG E ATTRIBUTI
- **CDATA (CHARACTER DATA)**
- ENTITÀ (*entity*)
- PI (*Processing Instructions*)
- COMMENTI

#### Sezione CDATA

<![CDATA[

]]>

Analizzata come XML (operazione di default)  
**(parsed)**

I caratteri rimangono nella loro forma originale  
**(unparsed)**

### XML – COMPONENTI DEL DOCUMENTO

- TAG E ATTRIBUTI
- CDATA (CHARACTER DATA)
- **ENTITÀ (ENTITY)**
- PI (*Processing Instructions*)
- COMMENTI

<rilegatore> Legatoria Libro &amp; Arte </rilegatore>

**Entità (entity)**  
**&amp;** ampersand

**Output: Legatoria Libro & Arte**

### XML – COMPONENTI DEL DOCUMENTO

- TAG E ATTRIBUTI
- CDATA (CHARACTER DATA)
- ENTITÀ (ENTITY)
- **PI (PROCESSING INSTRUCTIONS)**
- COMMENTI

`<?xml-stylesheet type="text/css" href="mySheet.css"?>`

**Processing Instructions**

Forma generale : `<?Finalità dell'istruzione?>`

- Il PI è un comando per il programma di elaborazione del documento XML per gestire il processo in modo preciso;
- I programmi che non riconoscono un determinato PI lo ignora semplicemente
- I documenti XML vengono in genere trattati da più programmi diversi.

<![CDATA

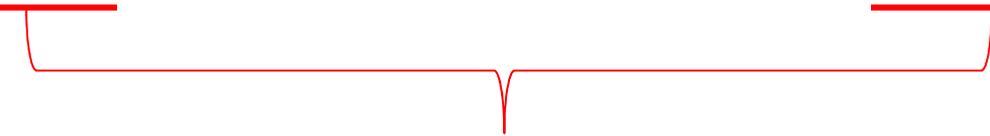
# XML: perché?

## XML: ortografia e sintassi

### XML – COMPONENTI DEL DOCUMENTO

- TAG E ATTRIBUTI
- CDATA (CHARACTER DATA)
- ENTITÀ (ENTITY)
- PI (PROCESSING INSTRUCTIONS)
- **COMMENTI**

`<!-- Testo di commento al programma o nota -->`



**Commento**, parte del file che non viene “processato”



«Un esperto è uno che conosce alcuni dei peggiori errori che può compiere nel suo campo, e sa come evitarli.»

*Werner Karl Heisenberg*

