

# Codifica di Testi - Introduzione XML Markup a.a. 2018-2019

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

*CNR-ILC-LicoLab*

Istituto di Linguistica Computazionale “A. Zampolli”,  
6th October 2018

# Contenuto della lezione

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare un documento XML e Definire uno Schema
  - Document Type Definition (DTD)
  - XML Schema Definition (XSD)
  - RELAX NG
- 4 Conclusioni

# Progress status

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni

# I linguaggi di codifica

## introduzione

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Definizione di codifica digitale del testo

Per **codifica** digitale dei testi intendiamo la *rappresentazione formale* di un **testo** ad un qualche livello descrittivo, su di un supporto digitale, in un formato utilizzabile da un elaboratore (*Machine Readable Form*) mediante un opportuno **linguaggio informatico** (F. Ciotti).

# I linguaggi di codifica

## Riassumendo

### Impostazione teorico-pratica

- un testo è molto di più della sequenza di caratteri che lo compongono
- per mezzo della codifica vogliamo rendere esplicite le caratteristiche che vogliamo analizzare
- solo quello che è esplicito può essere interpretato ed elaborato dal computer
- vogliamo codificare il testo per quello che è, non per quello che sembra
- codifica da effettuare mediante linguaggio di markup

# I linguaggi di codifica

## Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Il markup

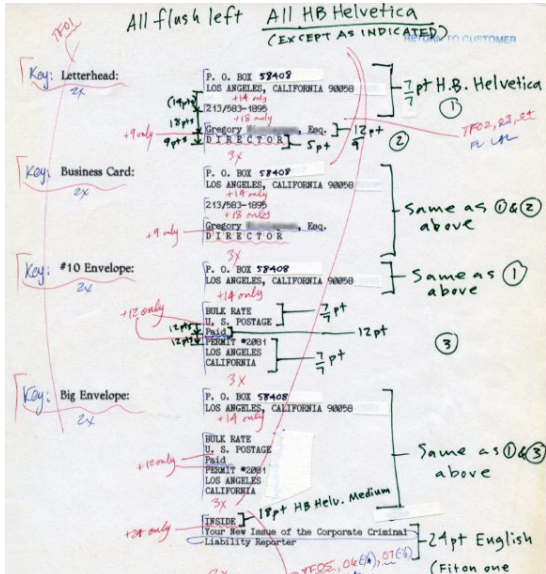
Il termine markup è stato utilizzato in passato per denotare i segni grafici che accompagnavano un testo apposti sul documento per indicare correzioni o modalità grafiche di stampa.

## Linguaggi di marcatura

2018-2019

A.M. Del  
Grosso

## I linguaggi di codifica



# I linguaggi di codifica

## Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso




















































I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

		Delete		Flush Left		Set in Bold Face Italic
		Insert		Flush Right		Set in Light Face
		Join		Center Horizontally		Wrong Font
		Move closer		Center Vertically		Hyphen
		Space		Move to the next line		En Dash
		Add Space		Move to the preceding line		Em Dash
		Delete Space		Indent 1 em		Superscript
		Transpose Word		Indent 2 ems		Subscript
		Transpose Letters		Paragraph		Comma
		To separate two or more marks		All Caps		Apostrophe
		Let it Stand (ignore correction)		Small Caps		Period
		Move Left		Caps & Small Caps		Semicolon
		Move Right		Capital Letter		Colon
		Move Up		Lower Case		Quotation Marks
		Move Down		Set in Roman		Parentheses
		Align Vertically		Set in Italic		Brackets
		Align Horizontally		Set in Bold Face		



# I linguaggi di codifica

## Linguaggi di marcatura

### Il markup

La codifica con linguaggi di marcatura (markup) è in sostanza un insieme di convenzioni, rese attraverso specifiche sequenze di caratteri, etichette, codici, (detti tags) intercalati nel testo per permettere agli elaboratori elettronici di distinguere le varie parti di un documento.

### Il markup formale

Un linguaggio di markup è un sistema formale per scambiare e pubblicare informazioni in formato testo in modo strutturato.

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

# I linguaggi di codifica

Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Il markup formale

Markup formale: costituito da un sistema ben preciso di istruzioni, ognuna delle quali è dotata di una specifica semantica e sintassi.

# I linguaggi di codifica

## Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Diversi tipi di markup

Esistono diversi linguaggi di markup, per rappresentare diversi tipi di documenti.

- Linguaggi procedurali (specific markup languages)
- Linguaggi dichiarativi (generic markup languages)

# I linguaggi di codifica

## Linguaggi di marcatura procedurale

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Linguaggi procedurali

- Orientati al documento, indicano come deve essere elaborato e disposto il testo
- Istruzioni da inserire nel testo per specificarne specifiche caratteristiche
- Font, dimensione, spaziatura del carattere, posizionamento nella pagina, colore, etc.

*Esempi: TeX e LaTeX, RTF*

# I linguaggi di codifica

## Linguaggi di marcatura procedurale

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Esempio RTF

```
{\rtf1\ansi\deff0\adeflang1025
{\fonttbl{\f0\froman\fprq2\fcharset0 Times New Roman;}
{\f1\froman\fprq2\fcharset0 Times New Roman;}
{\f2\fnil\fprq2\fcharset0 Lucida Sans Unicode;}
{\colortbl;\red0\green0\blue0;\red128\green128\blue128;}
{\stylesheet{\s1\cf0{\*\hyphen2\hyphlead2\hyphtrail2\hyphmax0}
\rtlch\af5\afs24\lang255\ltrch\dbch\af2\afs24\langfe255
\loch\f0\fs24\lang1040\snext1 Standard;}}
```

# I linguaggi di codifica

## Linguaggi di marcatura procedurale

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Esempio LaTeX

```
documentclass a4paper , 10pt article
```

# I linguaggi di codifica

Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Il markup procedurale

L'unico utilizzo di un testo codificato tramite un linguaggio procedurale è la creazione di un output orientato alla visualizzazione.

# I linguaggi di codifica

## Linguaggi di marcatura procedurale

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Scrivere la tesi di laurea con $\text{\LaTeX} 2_{\epsilon}$

Dipartimento di Ingegneria Meccanica,  
Nucleare e della Produzione  
Università di Pisa  
56126 Pisa PI

### Sommario

Lo scopo del presente articolo è fornire gli strumenti per scrivere una tesi di laurea utilizzando  $\text{\LaTeX} 2_{\epsilon}$ . Tale obiettivo è conseguito analizzando i problemi tipici incontrati durante la stesura della tesi e le possibili soluzioni; si pone particolare attenzione ai pacchetti da usare nelle varie circostanze. I singoli argomenti non vengono approfonditi nei dettagli ma si rimanda alla letteratura specifica o ad i manuali dei pacchetti suggeriti, ove necessario.

---

\*Ringrazio in primo luogo Fabiano Busdraghi che ha collaborato alla scrittura delle sezioni riguardanti le figure e gli oggetti flottanti. Ringrazio inoltre tutti coloro che mi hanno consigliato durante la stesura e la revisione di questo documento ed in particolare Claudio Beccari, Gustavo Cevolani, Massimo Guiggiani, Maurizio Himmelmann, Lorenzo Pantieri e Emiliano Vavassori.



# I linguaggi di codifica

## Linguaggi di marcatura dichiarativi

### Linguaggi dichiarativi

Orientati al testo, annotano la struttura, la funzione ed il significato degli elementi costitutivi del testo, tralasciandone l'aspetto.

- La posizione che il brano in questione occupa all'interno del documento (markup strutturale)
- Peculiarità del testo stesso (markup semantico)
- I fogli di stile definiscono la formattazione dell'output
- Molteplici usi del medesimo testo

*Esempio: famiglia SGML, XML*

# I linguaggi di codifica

## Linguaggi di marcatura dichiarativi

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Markup dichiarativi: contenuto e presentazione

La separazione tra contenuto e presentazione non solo è intenzionale, ma è la caratteristica principale di questi sistemi di marcatura: essa permette di concentrarsi sull'annotazione logica-semantica per funzioni di ricerca e di analisi, lasciando ad altro (ai fogli di stile) la resa grafica.

### Unico testo più usi

In questo modo si ha inoltre la possibilità di utilizzare uno stesso testo codificato con finalità o formattazioni differenti, a seconda delle varie esigenze.

# I linguaggi di codifica

## Markup dichiarativi: esempio SGML

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Standard Generalized Markup Language

```
1 <!DOCTYPE testo [  
2 <!ELEMENT testo (titolo?, paragrafo+)>  
3 <!ELEMENT titolo (#PCDATA)>  
4 <!ELEMENT paragrafo (#PCDATA)>  
5 ]>  
6 <testo>  
7   <titolo> Questo è il titolo del documento</titolo>  
8   <paragrafo> Questo è un paragrafo </paragrafo>  
9 </testo>  
10
```

# I linguaggi di codifica

## Markup dichiarativi vs Markup procedurali

### resa a video della frase

Le *Guidelines for Electronic Text Encoding and Interchange* sono *molto* complete e descrivono uno standard di *markup* del testo basato su XML.

Le `\textit{Guidelines for Electronic Text Encoding and Interchange}` sono `\textit{molto}` complete e descrivono uno standard di `\textit{markup}` del testo basato su XML.

`<titolo>`Le Guidelines for Electronic Text Encoding and Interchange`</titolo>` sono `<enfasi>`molto`</enfasi>` complete e descrivono uno standard di `<linguastraniera>` markup`</linguastraniera>` del testo basato su XML.

## LaTeX vs SGML

# I linguaggi di codifica

Linguaggi di marcatura

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## linguaggi semi-dichiarativi e/o semi-procedurali

Esistono anche linguaggi che possono essere definiti semi-procedurali, o semi-dichiarativi, che come si intuisce utilizzano le istruzioni sia per una codifica di tipo procedurale, sia per una codifica di tipo descrittivo o dichiarativo.

## HTML

HTML ha tra le sue etichette istruzioni di tipo procedurale per indicare come devono essere rese determinate porzioni di testo, e istruzioni di tipo dichiarativo che hanno una base semantica.

# Progress status

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

XXXXXXXXXXXXX

XML has its roots in Standard Generalized Markup Language (SGML), a language introduced in the 1980s that describes the structure and content of any machine-readable information.

XXXXXXXXXXXXXXXXXXXXx

XML can be thought of as a lightweight version of SGML. Like SGML, XML is a language used to create vocabularies for other markup languages

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

XXXXXXXXXXXXXXXXXXXX

It is a set of rules for defining custom-built markup languages

XXXXXXXXXXXXXXXXXXXX

XML was originally created to structure, store, and transport information. Like SGML, XML can be used to create XML applications or vocabularies, which are markup languages tailored to contain specific pieces of information.



# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

XXXXXXXXXXXXXXXXXXXX

XML is a markup language that is extensible

XXXXXXXXXXXXXXXXXXXX

developed and maintained by the World Wide Web Consortium (W3C), an organization created in 1994 to develop common protocols and standards for sharing information on the World Wide Web.

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

XXXXXXXXXXXXXXXXXX

XML, or eXtensible Markup Language, is a specification for storing information. It is also a specification for describing the structure of that information. And while XML is a markup language (just like HTML), XML has no tags of its own

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

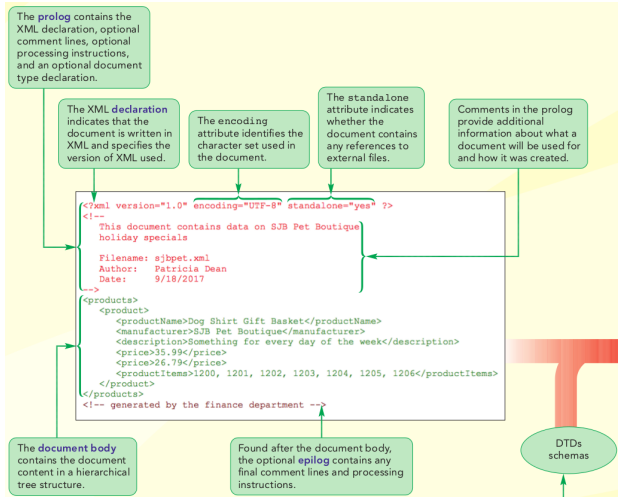


immagine dal libro *New Perspectives on XML, 3rd Edition*

# Fondamenti XML

## eXtensible Markup Language: Syntax Rules

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

- Every
- XML
- XML
- Every
- XML
- Some
- Comments

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## XML vista ad albero

XML is based on hierarchical trees in which order is significant.  
In XML, hierarchy and sequence are the main methods used to represent information.

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

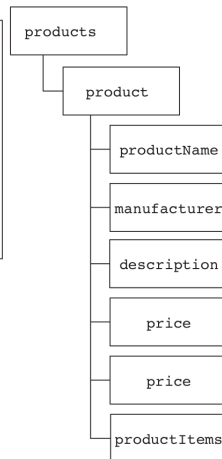
Conclusioni

XML document

```
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<!--
  This document contains data on SJB Pet Boutique
  holiday specials

  Filename: sjbpet.xml
  Author:   Patricia Dean
  Date:    9/18/2017
-->
<products>
  <product>
    <productName>Dog Shirt Gift Basket</productName>
    <manufacturer>SJB Pet Boutique</manufacturer>
    <description>Something for every day of the week</description>
    <price>35.99</price>
    <price>26.79</price>
    <productItems>1200, 1201, 1202, 1203, 1204, 1205, 1206</productItems>
  </product>
</products>
<!-- generated by the finance department -->
```

Hierarchy tree structure



# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

TEI-XML vocabulary

To meet the need of textual scholars,

XML Vocabularies

XML vocabularies

# Fondamenti XML

eXtensible Markup Language: Esempio TEI

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

```
<div type="narrative" n="6">
  <head>Sixth Narrative</head>
  <head>contributed by Sergeant Cuff</head>
  <div type="fragment" n="6.1">
    <opener>
      <dateline>
        <name type="place">Dorking, Surrey,</name>
        <date>July 30th, 1849</date>
      </dateline>
      <salute>To <name>Franklin Blake, Esq.</name> Sir, -</salute>
    </opener>
    <p>I beg to apologize for the delay that has occurred in the
      production of the Report, with which I engaged to furnish you.
      I have waited to make it a complete Report ...</p>
    <closer>
      <salute>I have the honour to remain, dear sir, your
        obedient servant </salute>
      <signed>
        <name>RICHARD CUFF</name> (late sergeant in the
          Detective Force, Scotland Yard, London). </signed>
      </closer>
    </div>
  </div>
```

*immagine dal sito TEI Guide Lines*



# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Manutenibilità

Because XML focuses on communicating the data, the overall structure is simple and easy to design and maintain.

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Documento ben formato (well-formed)

A well-formed document contains no syntax errors and satisfies the general specifications for XML code as laid out by the W3C. At a minimum, an XML document must be well formed or it will not be readable by programs that process XML code.

## Parti principali di un documento XML

An XML document consists of three parts—the prolog, the document body, and the epilog.

# Fondamenti XML

## eXtensible Markup Language: Esempio TEI

Codifica di

Testi -

Introduzione

XML Markup

a.a.

2018-2019

A.M. Del

Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following document is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονόσ' ἔτη. πᾶσι γὰρ δὴ πού τοις φιλολόγοις γνώριμα
              τὰ περὶ τῆς φιλοπονίας τάνδρὸς ἱστορούμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσαν τήνδε "κατέβην χθές
              εἰς Πειραιᾶ μετὰ Γλαύκωνος τοῦ Ἀριστάνου
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>

<!-- This document is not completed and was cut without a special meaning -->
```

## Parti principali di un documento XML

The prolog includes the following parts:

- XML declaration: indicates that the document is written in the XML language
- Processing instructions (optional): provide additional instructions to be run by programs that read the XML document
- Comment lines (optional): provide additional information about the document contents
- Document type declaration (DTD) (optional): provides information about the rules used in the XML document's vocabulary

# Fondamenti XML

eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Parti principali di un documento XML

The document body,

# Fondamenti XML

## eXtensible Markup Language: Prologo

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### XML declaration

```
<?xml version="version number" encoding="encoding  
type" standalone="yes|no" ?>
```

# Fondamenti XML

eXtensible Markup Language: Prologo

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## XML declaration: ERRORI

```
<?XML VERSION="1.0" ENCODING="ISO-8859-1"  
      STANDALONE="YES" ?>
```

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

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



# Fondamenti XML

## eXtensible Markup Language: Prologo

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### XML comments

comments are ignored by programs reading the document and do not affect the document's contents or structure.

### XML comments: sintassi

```
<!-- all'interno di questi segni il parser XML  
non entra -->
```

*METTERE TESTO COMMENTATO*

### mettere immagine XML-Prologo

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονῶς ἔτη. πᾶσι γὰρ δὴ που τοῖς φιλολόγοις γνῶριμα
              τὰ περὶ τῆς φιλοπονίας τάνδρὸς ἱστοροῦμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσιν τήνδε "κατέβην χθὲς
              εἰς Πειραιᾶ μετὰ Γλαῦκος τοῦ Ἀρίστωνος
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>
```

# Fondamenti XML

eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Esercizio prologo

Creare un file con estensione .xml ed inserire un prologo con la dichiarazione XML e un commento con le seguenti informazioni

## Esercizio prologo

```
<!--  
This document contains data on Corso di Codifica  
di Testi.  
Filename:  project.xml  
Author:   your name  
Date:    today's date  
-->
```

## XML parser

A program that reads and interprets an XML document is called an XML processor or XML parser, or simply a processor or parser.

## Cosa fa un XML parser

- Verifica che il documento rispetti le regole di sintassi XML (well-formed)
- Interpret PCDATA in a document and resolve any character or entity references found within the document
- Gestisce le processing instructions per interpretare eventuali dati

# Fondamenti XML

eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## XMLLINT

INTRODURRE XMLLINT Mettere immagine dell'uscita  
dell'help

## XML body

XML document is primarily composed of elements and attributes

The document body in an XML document is made up of elements that contain data.

Elements are the basic building blocks of XML files.

## Elementi: Sintassi

```
<element>content</element>
```

opening tag: `<element>;`

closing tag: `</element>`

## XML Element

*Gli elementi XML possono avere diversi tipi di contenuto:*

- contenuto strutturale: solo altri elementi, non testo
- contenuto misto: testo e anche altri elementi
- contenuto testuale: solo testo, non altri elementi

## XML Element: note importanti

- Element names are case sensitive
- Element names must begin with a letter or the underscore character
- Element names cannot begin with the string xml
- The name in an element closing tag must exactly match the name in the opening tag
- Element names can be used more than once (anche a diversi livelli gerarchici)
- Element names might be established already (es.: vocabolario TEI-XML)



# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### XML Element: empty e nested

- An open element or empty element is an element with no content (empty element).
- In addition to text content, elements also can contain other elements (nested element)

### XML esempi: empty e nested element

- `<element /> <element></element>`
- `<choice><sic>testo con errore</sic><cor> testo  
corretto</cor></choice>`

## XML Element: empty e nested

- An open element or empty element is an element with no content (empty element).
- In addition to text content, elements also can contain other elements (nested element)

## XML esempi: empty e nested element

- `<element /> <element></element>`
- `<choice><sic>testo con errore</sic><cor> testo  
corretto</cor></choice>`

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## XML Element: hierarchical relationship

- Nested
- Elements

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## mettere immagine parent - sibling

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονῶς ἔτη. πᾶσι γὰρ δὴ που τοῖς φιλολόγοις γνώριμα
              τὰ περὶ τῆς φιλοπονίας τάνορὸς ἱστοροῦμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσαν τήνδε "κατέβην χθές
              εἰς Πειραιᾶ μετὰ Γλαῦκωνος τοῦ Ἀρίστωνος
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>

<!-- This is not completed and was cut without a special meaning -->
```

## XML Element: hierarchical relationship

- All
- An
- hierarchy
- XML
- Syntax
- XML does

## XML Element: hierarchical relationship - Esercizio

Scrivere e fare il check di un xml non opportunamente annidato

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### XML Element: hierarchical relationship as tree structure

A

### mettere immagine xml-parent-child-quantifier

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονῶς ἔτη. πᾶσι γὰρ δὴ που τοῖς φιλολόγοις γνῶριμα
              τὰ περὶ τῆς φιλοπονίας τάνδρὸς ἱστοροῦμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσιν τήνδε "κατέβηναι χθὲς
              εἰς Πειραιᾶ μετὰ Γλαῦκωνος τοῦ Ἀρίστωνος
            </foreign>
          </p>
          ."
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>
```



### mettere immagine xml-parent-child-quantifier2

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονῶς ἔτη. πᾶσι γὰρ δὴ που τοῖς φιλολόγοις γνῶριμα
              τὰ περὶ τῆς φιλοπονίας τάνδρὸς ἱστοροῦμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσιν τήνδε "κατέβην χθές
              εἰς Πειραιᾶ μετὰ Γλαῦκωνος τοῦ Ἀρίστωνος
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>
```

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

XML Element: Mixed Content

XML Element: Mixed Content

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### XML Element: Esercizio

aprire il file XML non ben formato nella cartella xml del repository github e: - validarlo con XMILLINT - correggerlo (mettendo come commenti le correzioni fatte e brevemente spiegarle) – (nested, case sensitive) - aggiungere un figlio (child) all'elemento XYZ1 - aggiungere un fratello (sibling) all'elemento XYZ2

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## XML Attributi

Every element in an XML document can contain one or more attributes.

An attribute describes a feature or characteristic of an element

## XML Attributi

Attributes can only exist within an element. An attribute declaration really does not make sense without an element.

## XML Attributi

Attribute values are text strings. Therefore, an attribute value always must be enclosed within either single or double quotes.

## XML Attributi

```
<element attribute='value'> ... </element>  
    <element attribute='value' /> <element  
attribute='value', attribute2='value2' />
```

## XML Attributi: restrizioni ai nomi

- An attribute name must begin with a letter or an underscore ( \_ ).
- Spaces are not allowed in attribute names.
- Like an element name, an attribute name should not begin with the text string xml.

## XML Attributi

- attribute names are case sensitive
- The order of attributes is not significant

## XML Character and Entity References

- a numeric character reference syntax is `&#nnn;`;
- identified using a character entity reference by `&entity;`;

## XML References

- `&#61;`;
- `&amp;`;

## mettere immagine xml-Character-Entity

```
<?xml version="1.0" encoding="utf-8"?>
<?xml-stylesheet type="text/css" href="customStyle.css"?>
<!--The following is made online by the Perseus Project -->
<!--Added the TEI-lite DTD and a processing instruction -->
<!DOCTYPE TEI.2 SYSTEM "teixbaby.dtd">

<TEI.2>
  <text lang="en">
    <body>
      <div1 type="book" n="1" org="uniform" sample="complete">
        <div2 type="section" n="327A" org="uniform" sample="complete">
          <p>
            327A - 328B Socrates describes how he visited the Piraeus in company with Glauco, and
            was induced by Polemarchus and others to defer his return to Athens.
          </p>
          <p>
            <lemma lang="greek" targOrder="U" from="ROOT" to="DITTO">κατέβην κτλ.</lemma>
            Dionys. Hal.
            <title lang="la">de comp. verb.</title>
            p. 208 (Reiske)
            <foreign lang="greek">
              ὁ δὲ Πλάτων, τοὺς
              ἑαυτοῦ διαλόγους κτενίζων καὶ βοστρυχίζων, καὶ πάντα τρόπον ἀναπλέκων, οὐ
              διέλιπεν ὀγδοήκοντα γεγονῶς ἔτη. πᾶσι γὰρ δὴ που τοῖς φιλολόγοις γνῶριμα
              τὰ περὶ τῆς φιλοπονίας τάνορὸς ἱστοροῦμενα, τὰ τ' ἄλλα, καὶ δὴ καὶ τὰ
              περὶ τὴν δέλτον ἦν τελευτήσαντος αὐτοῦ λέγουσιν εὐρεθῆναι ποικίλως
              μετακειμένην τὴν ἀρχὴν τῆς πολιτείας ἔχουσαν τήνδε "κατέβην χθὲς
              εἰς Πειραιᾶ μετὰ Γλαῦκωνος τοῦ Ἀρίστωνος
            </foreign>
            ."
          </p>
        </div2>
      </div1>
    </body>
  </text>
</TEI.2>

<!-- This is not completed and was cut without a special meaning -->
```



## Text Character Parsing

text characters fall into three categories: parsed character data, character data, and white space.

## Text Character Parsing

- PCDATA
- CDATA
- White Space

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Parsed Character Data

Parsed character data (PCDATA) consists of all those characters that XML treats as parts of the code of an XML document.

## PCDATA

- XML declaration
- Opening tag e closing tag
- Character or entity references
- Commenti

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Parsed Character Data

The presence of PCDATA can cause unexpected errors to occur within a document

### XML PCDATA

## Character Data

Character data is not processed, but instead is treated as pure data content.

A CDATA section is a block of text that XML treats as character data only

## XML CDATA: sintassi

```
<![CDATA [ character data ]]>
```

## Character Data

You can use CDATA blocks when you want to include large blocks of special characters as character data

## CDATA: qualche vincolo

- You cannot use XML comments in a CDATA section.
- You cannot nest a CDATA section inside another CDATA section
- (cannot be nested within other CDATA sections).
- cannot be empty.

■ i simboli "]]" non sono ammessi perché questo è il marker

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

Esempio ed esercizio:

Inserire all'interno di un tag un frammento di codice HTML

CDATA: esempio

```
<htmlCode> <![CDATA[ <h1>Capitolo Primo</h1>  
  <h2>Sezione Seconda</h2> ]]> </htmlCode>
```

# Fondamenti XML

## eXtensible Markup Language

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## White Space

White space refers to nonprintable characters

## White Space: esempio

- White
- White
- White
- no

# Progress status

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)

XML Schema  
Definition (XSD)

RELAX NG

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

- Document Type Definition (DTD)
- XML Schema Definition (XSD)
- RELAX NG

4 Conclusioni



# Elementi per la definizione degli schemi xml

## principi

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

we need to make sure that the data that we receive follows a certain XML structure and should contain values which are coherent.

Your function needs to make sure that the caller passes correct XML data. You could make use of an XML Schema to perform this validation.

Performing such validations without the help of a SCHEMA will be extremely difficult most of the time.

# Elementi per la definizione degli schemi xml

## principi

Make sure that the XML document is structured exactly the way your function expects it to be.

We need an XML schema when we need to make sure that the XML document that we need to work with is in the expected format. Make sure that the values of elements and attributes are within the accepted range.

When data is managed and exchanged in XML format, there needs to be clear agreement about the structure of the XML document.

There needs to be a contract between the caller and the callee about the XML document being exchanged.

validate the XML document to make sure that it adheres to the format defined in the contract.

# Elementi per la definizione degli schemi xml

## principi

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

A Schema provides such the contract.  
It defines the structure of the XML document.  
It defines rules to validate the value of elements and attributes as well as their formats.  
Once a schema is defined, a Schema Validator can validate an XML document against the rules defined in the Schema.

# Elementi per la definizione degli schemi xml

## Tipi di formalismi per definire schemi XML

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

DTD, XDR, SOX, Schematron, DSD, DCD, DDML, RELAX  
NG

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

### Esempio DTD

An XML document may have a reference to an external DTD file or can have the DTD embedded as part of the XML file. The XML document given below has embedded DTD information.

```
<?xml version="1.0"?>
<!DOCTYPE Employee [
  <ELEMENT Name {
    <NAME> {
      <First> Middle, Last
    }
  }
  <ELEMENT First {#PCDATA}>
  <ELEMENT Middle {#PCDATA}>
  <ELEMENT Last {#PCDATA}>
]>
<Employee>
  <Name>
    <First>Jacob</First>
    <Middle>Sebastian</Middle>
    <Last>
      <Name>
        <First>
          <Middle>
            <Last>
              <Name>
                <Employee>
```

The example given below shows an XML document that refers to an external DTD file.

```
<?xml version="1.0"?>
<!DOCTYPE Employee SYSTEM "employee.dtd">
<Employee>
  <Name>
    <First>Jacob</First>
    <Middle>Sebastian</Middle>
    <Last>
      <Name>
        <First>
          <Middle>
            <Last>
              <Name>
                <Employee>
```

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Document Type Definition (DTD)

A document type definition describes the rule for an xml document structure, declaring the elements, attributes and entities that are part of the xml document

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Document Type Definition (DTD)

For an xml document to be valid, it must be well formed and also satisfies the document type definitions specified by the document type declaration.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Document Type Definition (DTD)

For an xml document to be valid, it must be well formed and also satisfies the document type definitions specified by the document type declaration.

well-formed document  $\neq$  valid document

any document that lacks a document type declaration may be well formed but cannot be valid.



# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Document Type Definition (DTD)

Document validation also aids file sharing since different applications that are aware of the generally agreed upon document type definition can produce xml document of similar structure and therefore easily communicate with each other through data exchange.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Document Type Definition (DTD)

included in the document' prolog

### Root element and content

declare the root element and its content model (children elements)

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Element Declaration (con figli)

```
<!ELEMENT element-name (child-element1,  
child-element-2 ...)>
```

### Element Declaration (solo testo)

```
<!ELEMENT element-name (#PCDATA)>
```

*Parsed character contents are plain texts with no child elements*

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Child Element Declaration

A child element declaration is done in the same way as the root element declaration using the `<!ELEMENT >` tag

### Element Declaration (root)

root element declaration must always come first.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition (DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Modificatori

Optional modifiers can also be used in element declarations to specify the number of times a child element may appear.

## Modificatori

- + One or more occurrences
- ? Zero or one occurrence
- \* Zero or more occurrences

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

### Modificatori

```
<!ELEMENT element-name (B, C)+ >
```

```
<!ELEMENT element-name (B+, C) >
```

```
<!ELEMENT element-name (B, C+) >
```

```
<!ELEMENT element-name (B+, C+) >
```

### Modificatori

If a child element must appear once then leave out the modifier.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Esercizio

- root: TEI - figli: header (obbligatorio una occorrenza);  
facsimile (opzionale una occorrenza); text (obbligatorio almeno  
una occorrenza) - header, facsimile, text hanno un content  
model testuale

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### choice declaration

```
<!ELEMENT element-name (child-a | child-b) >
```

### Dichiarazione di Choice

indicating a choice of just one, out of the list



# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Attributi

an attribute is declared with `<!ATTLIST >` element

### Attributi

An attribute is the property of an element that describes the element's content.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

### Attributi

```
<!ATTLIST Element-name Attr-name Attr-type  
Attr-state? default-value?>
```

### Attributi

“Element-name” is the name of the element “Attr-name” is the attribute to be declared. “Attr-type” specifies the expected attribute’s data type The “Attr-state” denota una tra i tre stati possibili di un attributo “default-value” if provided is the value to be used if the attribute is not supplied

# Elementi per la definizione degli schemi xml

## TABELLA dei tipi

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Tipi di attributi

# Elementi per la definizione degli schemi xml

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Stato di attributi

The state can be any of `#IMPLIED` (an optional attribute), `#REQUIRED` (a compulsory attribute) or `#FIXED` (a fixed value attribute that may not be changed) the value is provided as the default value. You cannot use the default-value with the `#REQUIRED` state as you must supply a value for the attribute.

# Elementi per la definizione degli schemi xml

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Mixed content - DTD

```
<!ELEMENT element-name (#PCDATA|child-element)* >
```

## Mixed content XML

```
<p>Ieri pomeriggio sono andato a  
<placeName>Pisa<placeName>, per un giro</p>
```

# Elementi per la definizione degli schemi xml

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Esercizio

root: TEI figli: header(obbligatorio una volta sola) -  
facsimile(opzionale una volta sola) - testo(obbligatorio una o  
più volte) - testo è un mixed content con possibile elemento  
<seg> attributi: - header: type:(fixed, CDATA “intestazione”);  
lang(opzional, NMTOKEN) - facsimile: source:(obbligatorio) -  
testo: id(obbligatorio, contenuto id) type(opzionale contenuto  
testuale)

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Empty elements

The only difference is that the child element name is substituted with the keyword “EMPTY”

### Empy content

```
<!ELEMENT element-name EMPTY>
```

### Empy content

```
<lb />
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Any elements

The content specification “ANY” when used in a declaration implies that any element as well as texts can be the child or content of the declared element.

### Any content

```
<!ELEMENT element-name ANY >
```



# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Dichiarare il tipo del documento XML

The document type declaration is placed in the xml document prolog; between the xml declaration and the root element.

### Dichiarare il tipo del documento XML

re-usability reasons, and then link it to the xml document and any other xml document through its URL

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## DOCTYPE

```
<!DOCTYPE root-element SYSTEM ‘‘External DTD’s  
URL’’ [Internal DTD ]>
```

```
<!DOCTYPE root-element [Internal DTD] >
```

```
<!DOCTYPE root-element SYSTEM ‘‘Ext-DTD URL’’ >
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Esercizi

- includere all'interno di un documento XML la dichiarazione del tipo e validare
- inserire nel prologo di un documento XML la dichiarazione del tipo di documento e validare.

*creare un file esterno con estensione .dtd prima di includerlo nel prologo XML.*

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Entity

Per includere dati da diverse fonti, DTD prevede l'uso di entità. Due tipologie di entità sono state definite: general entities e parameter entities.

### Entity: generiche e parametriche

- le general entities vengono espase nel documento XML
- le parameter entities vengono espase nel documento DTD

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Entity

- Le general entities si possono classificare in interne ed esterne; che a loro volta possono essere parsed oppure unparsed.
- Le parameter entities si possono classificare in interne ed esterne; che possono essere solo parsed.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)

XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

## Entity

Internal general entities help include special characters in an xml document, that would otherwise make an xml document become malformed if included literally.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Internal General Entity: Sintassi

```
<!ENTITY entity-name ‘‘replacement-string’’ or  
          ‘‘hexadecimal-code’’ >
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Internal General Entity: Sintassi

To use the entity name in the element content simply attach the ampersand to the beginning of the entity name and append the semi colon to the end like this; `&entity-name;` may include well formed xml elements.

### Internal General Entity: Sintassi

```
<!ENTITY firma ‘‘Angelo Mario Del Grosso’’>  
  
<p><salutation>&firma;<salutation></p>
```



# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Internal General Entity: Esempio con UNICODE

Spesso le entità vengono utilizzate per dare un nome ai riferimenti a carattere.

### Internal General Entity: Sintassi

```
<!ENTITY amaiuscola ‘‘&#65;’’>
```

```
<!ENTITY amaiuscola ‘‘&#x0041;’’>
```

```
<p><salutation>&amaiuscola;ddio<salutation></p>
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

### External General Entity

An xml document may be developed from other xml document located in different places. Grazie alle entità generiche esterne, questa eventualità diviene possibile.

### External General Entity: Sintassi

```
<!ENTITY entity-name SYSTEM ‘‘URL’’ >
```

```
<!ENTITY salutation SYSTEM ‘‘salut.xml.ent’’ >
```

*The URL points to the location of the external entity*

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

### External General Entity

External entities cannot contain a document type declaration as it will conflict with the main xml document type declaration. E' possibile utilizzare altre entità all'interno delle entità esterne.

### General Entity

though declared in the DTD, must be used in the document and not in the DTD itself.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Parameter Entity

Sono entità impiegabili all'interno del documento DTD. Ma non possono essere utilizzate all'interno del documento XML. There are two types of parameter entity; the internal and the external parameter entities.

### Parameter Entity: Sintassi

```
<!ENTITY % entity-name 'replacement-string'>
```

```
<!ENTITY % parameter-name SYSTEM 'URL' >
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Parameter Entity: Impiego

```
<!ENTITY % biblinfo ‘‘(title,author?,cost?)’’>  
  
<!ELEMENT biblInfo %biblinfo;>
```

### Parameter Entity: Impiego

```
<!ENTITY % biblInfo SYSTEM ‘‘biblInfo.dtd’’ >  
  
<!ELEMENT listBibl (bib+) >  
  
%biblInfo;
```

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Parameter Entity: utilità

When a parameter entity is inserted in the dtd, the parameter entity is replaced with the replacement content at execution time.

### Parameter Entity: utilità

This would make the dtd easier to develop and later maintained, if there are any changes to be made.

# Elementi per la definizione degli schemi xml

## Principi Document Type Definition

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

### Parameter Entity: utilità

External parameter entities enable the modularization and linking of document type definitions.

### Parameter Entity: utilità

with external parameter entities you can embed modularized document type definitions from different locations to form a single and more complete document type definition.

# Progress status

Codifica di  
Testi -  
Introduzione  
XML Markup  
a.a.  
2018-2019

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare un  
documento  
XML e  
Definire uno  
Schema

Document Type  
Definition  
(DTD)  
XML Schema  
Definition (XSD)  
RELAX NG

Conclusioni

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare un documento XML e Definire uno Schema

4 Conclusioni