

# 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”,  
4th October 2018

# Contenuto della lezione

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML
- 4 Document Typed Definition - DTD
- 5 XML Schema Definition - XSD
- 6 Conclusioni

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

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

Conclusioni

# Progress status

1 I linguaggi di codifica

2 Fondamenti del linguaggio XML

3 Validare XML

4 Document Typed Definition - DTD

5 XML Schema Definition - XSD

6 Conclusioni

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

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

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 XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

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

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

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

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 XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

Conclusioni

## Il markup

Il termine markup è stato utilizzato in passato per descrivere i segni grafici che accompagnavano un testo per indicare correzioni o come stampare un particolare passaggio.

## Nuove sfide

La discussione sulla rappresentazione digitale dei testi ha rivelato un intreccio di complesse questioni teoriche e pratiche.

# 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 XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

Conclusioni

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 Employee (Name)>
  <!ELEMENT Name (First, Middle, Last)>
  <!ELEMENT First (#PCDATA)>
  <!ELEMENT Middle (#PCDATA)>
  <!ELEMENT Last (#PCDATA)>
]>
<Employee>
  <Name>
    <First>Jacob</First>
    <Middle>V</Middle>
    <Last>Sebastian</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>V</Middle>
    <Last>Sebastian</Last>
  </Name>
</Employee>
```

# Progress status

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML
- 4 Document Typed Definition - DTD
- 5 XML Schema Definition - XSD
- 6 Conclusioni



# Progress status

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML**
- 4 Document Typed Definition - DTD
- 5 XML Schema Definition - XSD
- 6 Conclusioni

# Progress status

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML
- 4 Document Typed Definition - DTD**
- 5 XML Schema Definition - XSD
- 6 Conclusioni

# Progress status

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML
- 4 Document Typed Definition - DTD
- 5 XML Schema Definition - XSD**
- 6 Conclusioni

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

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

Conclusioni

# Progress status

- 1 I linguaggi di codifica
- 2 Fondamenti del linguaggio XML
- 3 Validare XML
- 4 Document Typed Definition - DTD
- 5 XML Schema Definition - XSD
- 6 Conclusioni**

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

A.M. Del  
Grosso

I linguaggi di  
codifica

Fondamenti  
del linguaggio  
XML

Validare XML

Document  
Typed  
Definition -  
DTD

XML Schema  
Definition -  
XSD

Conclusioni