

Codifica di Testi - XML Schema e Intro TEI a.a. 2018-2019

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

CNR-ILC-LicoLab

Istituto di Linguistica Computazionale “A. Zampolli”,
15th October 2018

Contenuto della lezione

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

1 XML Schema Definition (XSD)

2 Introduzione Text Encoding Initiative

3 Conclusioni

Progress status

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

1 XML Schema Definition (XSD)

2 Introduzione Text Encoding Initiative

3 Conclusioni

Elementi per la definizione degli schemi xml

principi XML Schema Definition

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Cos'è uno schema XML

Uno schema XML è un documento XML standard che descrive come deve essere realizzato un altro documento XML. Ci riferiamo a questa tecnologia con l'acronimo XSD.

A cosa serve uno Schema XML

I documenti XSD sono usati per validare documenti XML. Tuttavia un documento XSD viene realizzato tramite l'uso di un vocabolario predefinito riferibile attraverso un namespace con URI standard.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD Schema

Il termine XSD o XML Schema denota un documento XML che descrive e valida la struttura e il contenuto di un altro documento XML.

XSD Schema

Dichiarazione del documento (declaration) e istanza del documento (instance).

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD elemento root

L'elemento radice di uno schema XSD è sempre l'elemento "`<schema>`".

Tutte le definizioni devono seguire quindi l'elemento "`<schema>`".

XSD Schema

Tutti gli elementi e gli attributi dello schema sono dichiarati all'interno del namespace

"`http://www.w3.org/2001/XMLSchema.`".

Tutti i documenti XSD contengono la dichiarazione a questo namespace con prefisso convenzionale **xsd** oppure **xs**.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD componenti di base

I componenti di base di uno Schema XSD sono le dichiarazioni degli elementi e le dichiarazioni degli attributi.

XSD Schema

Le dichiarazioni più complesse si poggiano su queste unità: elementi e attributi.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD dichiarazioni

Scrivere un pezzo di codice XSD per descrivere e validare un elemento per un documento XML è detto *element declaration*.

XSD dichiarazioni di base

XSD permette di dichiarare elementi, attributi e di specificare il numero di figli, le occorrenze, l'ordine di apparizione, e i tipi di dati del content model.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element Types: simple and complex

La dichiarazione di un elemento può avere un tipo semplice (*simple type*) oppure un tipo complesso (*complex type*) a seconda della sua struttura e del suo contenuto.

Simple Type e Complex Type

La dichiarazione di un elemento ha un tipo semplice se non possiede **né figli né attributi**.

La dichiarazione di un elemento ha un tipo complesso in tutti gli altri casi.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD esempio

```
<xsd:schema  
  xmlns:xsd='http://www.w3.org/2001/XMLSchema''>  
  <xsd:element name='text' />  
</xsd:schema>
```

XSD esempio elemento di tipo semplice

```
<text>Il primo documento XML Validato</text>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XML XSD esempio

Il documento XML istanza dello schema XSD per essere valido deve contenere un elemento radice. Validare il documento XML con il relativo XSD con XMLlint.

XMLlint

```
xmllint xmlfirst.xml --schema  
../schema/xsd/xsdfirst.xsd
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element Complex Types: esempio

```
<xsd:schema
xmlns:xsd='http://www.w3.org/2001/XMLSchema''>
<xsd:element name='Employee'> <xsd:complexType>
<xsd:attribute name='FirstName'/>
</xsd:complexType> </xsd:element> </xsd:schema>
```

Element Complex Types: esempio

Il documento XML istanza dello schema:

```
<Employee FirstName="Jacob"/>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Types

Alla base dello standard XSD ci sono le dichiarazioni degli elementi e degli attributi, ad un livello di astrazione più alto ci sono i types e i groups.

Element Complex Types

Un tipo complesso (Complex Type) può avere attributi e/o elementi figli.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element Complex Types: Esempio Elemento con Figlio

```
<xsd:schema
  xmlns:xsd='http://www.w3.org/2001/XMLSchema'>
  <xsd:element name='text'> <xsd:complexType>
    <xsd:sequence> <xsd:element name='body'>/>
  </xsd:sequence> </xsd:complexType> </xsd:element>
</xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Espressività dell'XSD

- Attribute Group, Element Group
- Order Indicators: all, sequence, choice
- Occurrence Indicators: minOccurs and maxOccurs
- Annotation (utili per documentare le dichiarazioni)

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Espressività dell'XSD

- Data types: Built-in
- FACETS per una validazione oculata dei valori (elemento o un attributo).
- Restriction on base attribute
- Numeric e String Data Type
- Pattern facet

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element Declaration

Una dichiarazione di elemento (**XSD element declaration**) rappresenta un elemento nel documento XML istanza dello schema.

Element Declaration: Istanza XML

`<p>il contenuto testuale di un paragrafo</p>`

*Un elemento ha un tipo semplice (**simple type**) se non ha attributi e non ha elementi figli*

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Global and Local element declaration

Il vantaggio di usare dichiarazioni di elementi globali (**global element declaration**) è che la dichiarazione può essere riferita più volte da altre dichiarazioni all'interno del documento XSD.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Global element declaration: esempio

```
<xsd:schema
xmlns:xsd='http://www.w3.org/2001/XMLSchema">
  <xsd:element name='body'> <xsd:complexType>
    <xsd:attribute name='lang'/'> <xsd:attribute
name='type'/'> </xsd:complexType> </xsd:element>
  <xsd:element name='text'> <xsd:complexType>
    <xsd:all> <xsd:element ref='body'/'> </xsd:all>
  </xsd:complexType> </xsd:element> </xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Group declaration

Attributi e elementi possono essere raggruppati in costrutti detti **Attribute Groups** e **Element Groups**.

E' possibile riferirsi a questi gruppi all'interno di altre dichiarazioni ottenendo un significativo grado di riusabilità.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Global element declaration: esempio

```
<xsd:group name='fileDesc'> <xsd:sequence>  
<xsd:element name='titleStmt'></xsd:element>  
name='publicationStmt'></xsd:element>  
name='sourceDesc'></xsd:sequence>
```

Global element declaration: esempio

```
<xsd:group ref='fileDesc'>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element Declaration: Attributi nella dichiarazione

L'unico attributo *obbligatorio* nella dichiarazione di elementi è l'attributo **name**.

Element Declaration: Attributi - Esempio

```
<xsd:element name='TEI' />
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element declaration: lista Attributi

- **name** (*g-l*): il nome dell'elemento, unico attributo obbligatorio (*dare nomi significativi agli elementi*)
- **type** (*g-l*): aggiunta di un tipo di dato (*data type*) (XSD ha built-in data types)
- **id** (*g-l*): identificatore univoco della dichiarazione dell'elemento (non ha effetti sull'istanza XML)
- **default** (*g-l*): specifica il valore di default di un elemento (solo se l'elemento è opzionale)

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element declaration: lista Attributi (cont.)

- **fixed** (*g-l*): valore di un elemento predefinito (“fixed” e “default” sono mutuamente esclusivi)
- **nilable** (*g-l*): indica che l’elemento è vuoto (xsd:nil nell’istanza XML del documento)
- **block** (*g-l*): è utilizzato per controllare la sostituzione degli elementi (può assumere 4 differenti valori: *substitution*, *extension* *restriction* e *#all*)

Elementi per la definizione degli schemi xml

principi XSD

Element declaration: lista Attributi (cont.)

- **final** (*g*): limita la dichiarazione dei gruppi di sostituzione (può assumere 3 valori: *#all*, *restriction* and *extension*)
- **abstract** (*g*): un elemento astratto non può essere istanziato (è un modo per forzare la sostituzione)
- **substitutionGroup** (*g*): è un modo per avere grande estensibilità delle dichiarazioni (facilita la creazione di nuovi tipi di elementi derivati)
- **minOccurs** (*l*): controlla il numero di occorrenze minime degli elementi all'interno del documento XML istanza dello schema. (**minOccurs=0** indica elemento *optionale*).

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element declaration: lista Attributi (cont.)

- **maxOccurs** (*l*): controlla il numero massimo di occorrenze degli elementi all'interno del documento XML istanza dello schema
- **ref** (*l*): riferimento ad altri elementi dichiarati globalmente all'interno dello schema (riusabilità)
- **form** (*l*): specifica se l'elemento deve essere qualificato con l'aggiunta del namespace ("qualified" oppure "unqualified").

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Element declaration: Attributi - Esempio

```
<xsd:element name='body'> <xsd:complexType>  
  <xsd:sequence maxOccurs='unbounded'>  
    <xsd:element ref='div' /> </xsd:sequence>  
  </xsd:complexType> </xsd:element>
```

Element declaration: Attributi - Esempio

```
<xsd:element name='div' type='divType' />  
<xsd:complexType name='divType'> [...]  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration

Un attributo è dichiarato utilizzando l'elemento
`<xsd:attribute>`.

L'unico attributo obbligatorio nella dichiarazione è **"name"**.

Attribute declaration: Global vs Local

Quando una dichiarazione di attributo ha il livello gerarchico successivo a `"<xsd:schema>"` (*la root dello schema*), questa dichiarazione è detta globale (*global attribute declaration*). In caso contrario l'attributo è all'interno di un *Complex Type*, ed è una (*Local attribute declaration*).

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration

```
<xsd:attribute name='Name'  
type='xsd:string' />
```

Qualsiasi dichiarazione di attributo per essere effettivamente utile deve essere riportata all'interno di una dichiarazione di elemento

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Attributi

- **name** (*g-l*): il nome dell'attributo, attributo obbligatorio.
- **id** (*g-l*): l'identificativo della dichiarazione all'interno del documento XSD (non influenza l'istanza XML).
- **type** (*g-l*): associa un tipo di dato allo spazio di valori che può assumere l'attributo.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Attributi (cont.)

- **default** (*g-l*): assegna un valore di default all'attributo (solo se l'attributo non è presente).
- **fixed** (*g-l*): assegna un valore predefinito e fisso all'attributo.
- **ref** (*l*): riferimento ad una dichiarazione di attributo globale

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Attributi (cont.)

- **use** (*l*): specifica se l'attributo è opzionale o meno.
Possibili valori sono *optional*, *required*, *prohibited*
- **form** (*l*): indica se l'attributo deve avere o meno indicato il namespace all'interno del documento XML istanza dello schema (può assumere due valori: *qualified* e *unqualified*)

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Esempio (global)

```
<xsd:schema
  xmlns:xsd='http://www.w3.org/2001/XMLSchema'>
  <xsd:attribute name='analysis' /> <!-- -->
  <xsd:element name='word'> <xsd:complexType>
    <xsd:attribute ref='analysis' />
  </xsd:complexType> </xsd:element> </xsd:schema>
```

Global attribute declarations are useful when an attribute is declared with several validations

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Attribute Groups

Definire gruppi di attributi permette di riusare le dichiarazioni degli attributi in molteplici altre dichiarazioni all'interno dello schema.

I gruppi di attributi permettono di avere un maggiore livello di riusabilità delle dichiarazioni.

Attraverso l'uso dei gruppi di attributi si evitano inutili ripetizioni nel codice XSD ed è possibile costruire catene di gruppi strutturate in gerarchie ben controllate

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Esercizio (global)

```
<xsd:attributeGroup name='EmpAttributes'>
  <xsd:attribute name='name'> <xsd:simpleType>
    <xsd:restriction base='xsd:string'>
      <xsd:maxLength value='20'> </xsd:restriction>
    </xsd:simpleType> </xsd:attribute> <xsd:attribute
    name='department'> <xsd:simpleType>
      <xsd:restriction base='xsd:string'> <xsd:length
      value='2'> </xsd:restriction> </xsd:simpleType>
    </xsd:attribute> </xsd:attributeGroup>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Attribute declaration: Esercizio (global)

```
<xsd:element name='Employees'>
  <xsd:complexType> <xsd:sequence> <xsd:element
name='Manager'> <xsd:complexType>
  <xsd:attributeGroup ref='EmpAttributes' />
</xsd:complexType> </xsd:element> <xsd:element
name='TechLead'> <xsd:complexType>
  <xsd:attributeGroup ref='EmpAttributes' />
</xsd:complexType> </xsd:element> </xsd:sequence>
</xsd:complexType> </xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD Data Type

I linguaggi di programmazione uso i tipi (*data types*) per controllare i valori che può assumere una variabile e le operazioni che possono essere effettuate su quest'ultime.

In parole povere, quando si associa un data type ad una variabile si vuole restringere lo spazio di valori permesso e restringere le operazioni che possono elaborare tali valori.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD Data Type: esempio - dichiarazione

```
<xsd:element name='div'> <xsd:complexType>  
  <xsd:attribute name='type' /> <xsd:attribute  
name='n' type='xsd:integer' />  
</xsd:complexType> </xsd:element>
```

XSD Data Type: esempio - istanza

```
<div type="chapter" n="1" />' (corretto)  
<div type="chapter" n="uno" />' (errato)
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD Data Type

XSD supporta diversi tipi di dati, per descrivere e validare quasi tutti i valori che possono occorrere in un documento XML.

Ma supporta anche la possibilità di derivare nuovi data types a partire da quelli disponibili (built-in

XSD Data Type

I tipi di dati aiutano a descrivere un certo dato in modo più accurato, contribuendo a validarne i valori in modo più efficace.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

XSD Data Type

XSD supporta una cinquantina di tipi di dato divisi in due categorie: Primitive Data Types e Derived Data Types.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Primitive Data Types

Primitive Data Types sono tipi di dati dai quali sono stati derivati altri tipi di dati.

Primitive Data Types

string	boolean	decimal
float	double	duration
dateTime	time	date
gYearMonth	gYear	gMonthDay
gDay	gMonth	hexBinary
base64Binary	anyURI	QName
NOTATION		

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Primitive Data Types

Primitive Data Types non derivano da alcun tipo di base.

Derived Data Types

Nel sistema di tipi di XSD esistono data types che sono direttamente oppure indirettamente derivati dai Primitive Data Types.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Derived Type	Base Type
normalizedString	string
Integer	decimal

Derived Type	Base Type
Token	normalizedString
NMTOKEN	token
Name	token
Language	token
NMTOKENS	NMTOKEN
NCName	Name
ENTITY	NCName
ENTITIES	ENTITY

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Derived Type	Base Type
nonPositiveInteger	integer
nonNegativeInteger	integer
Long	integer
negativeInteger	nonPositiveInteger
int	Long
short	int
byte	short
unsignedLong	nonNegativeInteger
positiveInteger	nonNegativeInteger
unsignedInt	unsignedLong
unsignedShort	unsignedInt
unsignedByte	unsignedShort

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Data Type: Facets

Ciascun tipo di dato a una collezione di proprietà che possono essere usate per attuare validazioni aggiuntive ai valori permessi dal tipo di dato corrente.

*Queste proprietà sono chiamate **Facets** in XSD.*

Data Type: Facets

Ogni tipo ha un insieme stabilito di Facets che controllano una certa proprietà o una certa caratteristica del tipo di dato considerato.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Data Type: Facets Esempio

```
<xsd:element name='name'> <xsd:complexType>  
<xsd:attribute name='type'> <xsd:simpleType>  
<xsd:restriction base='xsd:string'> <xsd:length  
value='15'> </xsd:restriction>  
</xsd:simpleType> </xsd:attribute>  
</xsd:complexType> </xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Data Type: Facets Esempio

```
<xsd:element name='name'> <xsd:complexType>  
<xsd:attribute name='type'> <xsd:simpleType>  
  <xsd:restriction base='xsd:string'> <xsd:patter  
    value=''[A-Za-z]+''/> </xsd:restriction>  
</xsd:simpleType> </xsd:attribute>  
</xsd:complexType> </xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

	pattern	whitespace	length	minLength	maxLength	enumeration	totalDigits	fractionDigits	minInclusive	maxInclusive	minExclusive	maxExclusive
string	✓	✓	✓	✓	✓	✓						
boolean	✓	✓										
decimal	✓	✓				✓	✓	✓	✓	✓	✓	✓
float	✓	✓				✓			✓	✓	✓	✓
double	✓	✓				✓			✓	✓	✓	✓
duration	✓	✓				✓			✓	✓	✓	✓
dateTime	✓	✓				✓			✓	✓	✓	✓
time	✓	✓				✓			✓	✓	✓	✓
date	✓	✓				✓			✓	✓	✓	✓
gYear	✓	✓				✓			✓	✓	✓	✓
gMonthDay	✓	✓				✓			✓	✓	✓	✓
gDay	✓	✓				✓			✓	✓	✓	✓
gMonth	✓	✓				✓			✓	✓	✓	✓
base64Binary	✓	✓	✓	✓	✓	✓						
base64Binary	✓	✓	✓	✓	✓	✓						
anyURI	✓	✓	✓	✓	✓	✓						
QName	✓	✓	✓	✓	✓	✓						

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Simple Type vs Complex Type

La differenza sostanziale tra un tipo semplice (*simple types*) e un tipo complesso (*complex types*) è che solo un tipo complesso può avere elementi figli e attributi.

Simple Type vs Complex Type

Le entità *simple type* possono trattare solo valori destrutturati. Sia gli elementi sia gli attributi possono essere *simple type*.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="ZipCode" type="zipType" />
  <xsd:simpleType name="zipType">
    <xsd:restriction base="xsd:integer">
      <xsd:maxInclusive value="99999"/>
      <xsd:minInclusive value="10000"/>
    </xsd:restriction>
  </xsd:simpleType>
</xsd:schema>
```

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <xsd:element name="ZipCode">
    <xsd:simpleType>
      <xsd:restriction base="xsd:integer">
        <xsd:maxInclusive value="99999"/>
        <xsd:minInclusive value="10000"/>
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>
</xsd:schema>
```

*Simple Types can be declared **globally** or **locally***

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Global simple types

le definizioni di tipi globali sono utili per il riuso del codice e per mantenere ed organizzare al meglio una schema XSD ampio.

Molto utili quando bisogna uniformare un insieme di validazioni e mantenerle in modo coerente

Elementi per la definizione degli schemi xml

principi XSD

simple types: esercizio

```
<xsd:simpleType name='chapterNumberType'>  
<xsd:restriction base='xsd:integer'>  
<xsd:maxInclusive value='1000'>  
<xsd:minInclusive value='1'> </xsd:restriction>  
</xsd:simpleType>
```

simple types: esercizio

```
<xsd:element name='item'> <xsd:complexType>  
<xsd:attribute name='originalChapter'  
type='chapterNumberType'> </xsd:complexType>  
</xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving

Un nuovo tipo può essere derivato da un tipo già dichiarato (primitivo o meno) ed ereditarne le caratteristiche.

simple types example

- Derive by restriction
- Derive by list
- Derive by Union

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving

Identificare il **tipo di dato di base** e aggiungere le dichiarazioni di **restrizione** e le **regole di validazione** che si reputano utili.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving by Restriction

A restriction is defined by adding Un restrizione si definisce aggiungendo “**xsd:restriction**” alla dichiarazione del *Simple Type*.

Ciascun tipo di dato possiede una collezione di proprietà (facets) attraverso le quali definire la restrizione voluta.

Elementi per la definizione degli schemi xml

principi XSD

simple types: deriving by Restriction

```
<xsd:simpleType name='signatureType'>  
  <xsd:restriction base='xsd:integer'>  
    <xsd:totalDigits value='5'></xsd:restriction>  
  </xsd:simpleType>
```

simple types: deriving by Restriction

```
<xsd:element name='signature'  
type='signatureType'>  
  
  <signature>12345</signature>  (valido)  
  <signature>123ab</signature>  (non valido)
```


Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving by List

Un data type può contenere una lista di valori separata da spazi

simple types: deriving by List

```
<xsd:simpleType name='chapterNumberList'>  
  <xsd:list itemType='xsd:integer' />  
</xsd:simpleType>  
  <xsd:element name='chapters'  
type='chapterNumberList' />  
  <chapters>1 53 60 61 205 409</chapters>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving by Union

Un tipo derivato può accettare valori di qualsiasi tipo dichiarato nella derivazione.

simple types: deriving by Union

```
<xsd:simpleType name='ZipCityUnion'> <xsd:union>  
<xsd:simpleType> <xsd:restriction base="ZipType"/>  
</xsd:simpleType> <xsd:simpleType>  
<xsd:restriction base="CityType"/>  
</xsd:simpleType> </xsd:union> </xsd:simpleType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving

I valori sono accettati solo se sono validi per uno dei tipi dichiarati nella derivazione.

simple types: deriving

Nei costrutti di derivazione *non è permesso aumentare lo spazio dei valori consentiti*.

Non è possibile quindi definire regole meno restrittive di quelle che caratterizzano il tipo di base.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

simple types: deriving facets

Quando si deriva un tipo semplice da un altro tipo semplice, tutte le facets vengono ereditate.

simple types: deriving facets

E' possibile utilizzare l'attributo **fixed** di una data proprietà (facets) a **true** per assicurare che la proprietà stessa non venga modificata dal tipo derivato.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Simple types: controllare la derivazione

XSD ha anche un meccanismo per controllare e proteggere (vietare) un tipo dall'essere derivato.

Si impiega l'attributo **final** della dichiarazione di simple type.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

controllare la derivazione: l'attributo final

- restriction
- list
- union
- extension
- #all

Elementi per la definizione degli schemi xml

principi XSD

Simple types: controllare la derivazione - esempio

```
<xsd:schema
  xmlns:xsd='http://www.w3.org/2001/XMLSchema''>
  <xsd:simpleType name='zipType'
    final='restriction union list extension'>
    <xsd:restriction base='xsd:integer'>
      <xsd:maxInclusive value='99999' fixed='true'/>
      <xsd:minInclusive value='10000'/>
    </xsd:restriction>
  </xsd:simpleType>
</xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Simple types: controllare la derivazione - esempio

Il termine estensione si riferisce alla possibilità di derivare un nuovo tipo che risulta essere complex type da un tipo base simple type.

Simple types: controllare la derivazione - esempio

Quando l'attributo **final** ha valore **#all**, il Simple Type non può essere derivato in alcun modo.

Elementi per la definizione degli schemi xml

principi XSD

Complex Type

Un tipo coompleso (**complex type**) può essere strutturato e quindi avere figli e avere attributi.

Named Complex types

Una complex type può avere un nome (*Named Complex Type*) così da poter essere impiegato all'interno della dichiarazione di altri tipi complessi.

le Attribute declarations non possono avere Complex Types.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Named Complex Type Esempio

```
<xsd:complexType name='addressType'> <xsd:all>  
<xsd:element name="address"/> <xsd:element  
name="street"/> <xsd:element name="settlement"/>  
<xsd:element name="country"/> </xsd:all>  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Named Complex types Esempio

```
<xsd:element name='AddressDivision'>  
  <xsd:complexType>  
    <xsd:sequence>  
      <xsd:element  
        name="addressee"/>  
      <xsd:element name="address"  
        type="addressType"/>  
    </xsd:sequence>  
  </xsd:complexType>  
</xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type

I Named Complex Types offrono un potente strumento per riusare parti di codifica XSD all'interno dello schema.

Mentre Anonymous Complex Types non possono essere riferiti e appaiono solo all'interno della dichiarazione in cui vengono definiti.

Anonymous Complex types

```
<xsd:element name='Customer'> <xsd:complexType>  
<xsd:sequence>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Content Model

La struttura di un Complex Type è chiamata Content Model.

Content Model: Simple Content, Complex Content

- **Simple Content:** contiene valori testuali e può avere attributi.
- **Complex Content:** *empty, element- only , mixed type.*

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Content Model: complex content

- *Empty* content può avere attributi.
- *element-only* ha elementi figli, può avere attributi, ma non può avere testo come valore dell'elemento.
- *mixed content* può contenere elementi figli, attributi e valori testuali.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: Simple Content

Simple Content: text value e attributi, ma non può avere figli.

Complex type: Simple Content esempio

```
<xsd:element name='name'> <xsd:complexType>  
<xsd:simpleContent> <xsd:extension  
base='xsd:string'> <xsd:attribute name="type"  
type="xsd:string"/> </xsd:extension>  
</xsd:simpleContent> </xsd:complexType>  
</xsd:element>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: Complex Content - Empty content

Gli elementi Complex Type che hanno un modello del contenuto complesso vuoto sono simili agli elementi di tipo complesso con contenuto semplice.

Complex type: Empty content

Gli elementi Empty content complex types non possono contenere testo, ma possono avere attributi.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: element-only content

Gli elementi Complex Types element-only content hanno un modello del contenuto composto solo da elementi figli e attributi, ma non possono contenere del testo.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: element-only content - esempio

```
<xsd:group name='fileDesc'> <xsd:sequence>  
<xsd:element name="titleStmt"/> <xsd:element  
name="publicationStmt"/> <xsd:element  
name="sourceDesc" /> </xsd:sequence> </xsd:group>  
<xsd:element name='teiHeader'> <xsd:complexType>  
<xsd:sequence> <xsd:group ref='fileDesc' />  
</xsd:sequence> </xsd:complexType> </xsd:element>  
</xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: mixed content

Gli elementi Mixed Content hanno un modello del contenuto che combina il simple content e l'element-only content.

Può contenere elementi figli, testo e attributi.

Complex type: mixed content

Un parser XML può estrarre molte più informazioni se un elemento ha un modello del contenuto The Mixed content.

Elementi per la definizione degli schemi xml

principi XSD

Complex type: mixed content - esempio

```
<xsd:element name='seg'> <xsd:complexType  
mixed='true'> <xsd:all> <xsd:element  
name="name"/> <xsd:element name="quote"/>  
</xsd:all> </xsd:complexType> </xsd:element>  
</xsd:schema>
```

Complex type: mixed content - esempio

```
<seg><name>Petrarca</name> disse:  
<quote>Raramente la grande bellezza e la grande  
virtù dimorano assieme</quote></seg>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex type: mixed content

*La sola differenza con il modello di contenuto element-only è la presenza dell'attributo “**mixed**.” Nel campo delle Digital Scholarly Editions ci sono moltissimi casi in cui utilizziamo questo tipo di contenuto misto*

Elementi per la definizione degli schemi xml

principi XSD

Complex type con attributi

Se un elementoo di tipo complesso contiene dichiarazioni di elementi e di attributi, la dichiarazione degli attributi **segue** la dichiarazione degli elementi.

Complex type: mixed content

```
<xsd:element name='seg'> <xsd:complexType  
  mixed='true'> <xsd:sequence> <xsd:element  
  name="quote"/> </xsd:sequence> <xsd:attribute  
  name="part"/> </xsd:complexType> </xsd:element>  
</xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Elementi e Attributi: ordering

Gli attributi di un elemento possono apparire in qualsiasi ordine.
La posizione (ordine) di apparizione non è significativa.

Elementi e Attributi: ordering

La posizione (ordine) degli elementi all'interno di un documento XML è significativa. Quindi è importante avere un meccanismo per specificare la **politica di ordinamento** all'interno del content model di un elemento.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Elementi e Attributi: ordering

- **sequence**: è un elemento (`<xsd:sequence />`) che indica che gli elementi dichiarati devono seguire esattamente l'ordine specificato.
- **all**: è un elemento (`<xsd:all>`) che indica che l'ordine degli elementi non è significativo.
- **choice**: è un elemento (`<xsd:choice />`) che indica che solo uno dell'insieme di elementi specificato può essere usato nel documento XML.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Elementi e Attributi: occurrence indicator

Gli attributi non possono apparire più di una volta all'interno dell'elemento.

Elementi e Attributi: occurrence indicator

Un Elemento può apparire più di una volta all'interno del proprio elemento padre/contenitore.

Possiamo controllare quante occorrenze di un elemento sono consentite grazie agli attributi **minOccurs** e **maxOccurs**

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

minOccurs	maxOccurs	Result
1	1	The element is mandatory and should appear only once.
0	1	The element is optional. It may appear once but not more than once.
0	5	The element is optional and can appear up to 5 times.
1	5	The element is mandatory and can appear 5 times maximum.
0	unbounded	The element is optional and can appear any number of times.
1	unbounded	The element is mandatory and can appear any number of times
2	2	The element should appear exactly 2 times.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Content Model: Deriving

Nuovi tipi complessi possono essere derivati da tipi complessi con modello del contenuto Simple Content. *Non è possibile derivare Complex Content da Simple Content.*

Content Model: Complex Content

Un elemento con modello Complex Content può contenere elementi figli, attributi e testo.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Content Model: Complex Content

Nuovi complex types possono essere derivati da complex type già definiti by restriction oppure by extension

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: da simple type

E' possibile derivare complex type da un tipo base simple type, ma solo se l'elemento è simple content.

Se l'attributo *final* è impostato sul valore **extension**, il tipo base (simple type) non può essere esteso.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: da simple type - esempio

```
<xsd:complexType name='divTypeEx'>  
  <xsd:simpleContent> <xsd:extension  
base='divType'> <xsd:attribute name='type',  
use='required' /> </xsd:extension>  
</xsd:simpleContent> </xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: Diversi casi

- derivare un nuovo tipo by restriction oppure by extension da un complex type simple content model
- derivare by restriction da un complex type con simple content
- aggiungere restrizioni al content/text di un elemento
- aggiungere restrizioni ad un attributo
- eliminare uno o più attributi ad un tipo di elemento

Le politiche di derivazione cambiano a seconda del modello di contenuto del tipo considerato

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation: esempio

```
<xsd:complexType name='RestrictedPhoneType'>
  <xsd:simpleContent> <xsd:restriction
    base='PhoneType'> <xsd:pattern
      value="[0-9]3-[0-9]3-[0-9]4"/> <xsd:attribute
        name='Type'> <xsd:simpleType> <xsd:restriction
          base='xsd:string'> <xsd:enumeration
            value="Home"/> <xsd:enumeration value="Work"/>
        </xsd:restriction> </xsd:simpleType>
        </xsd:attribute> <xsd:attribute
          name="CallOnWeekend" type="xsd:boolean"
          use="prohibited"/> </xsd:restriction>
      </xsd:simpleContent> </xsd:complexType>
```


Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation

Si noti l'attributo **use** con valore *prohibited*.

Se si vuole eliminare un attributo è necessario ridefinirlo nel tipo derivato e impostare l'attributo *use* al valore *prohibited*.

Complex Type derivation

Se un attributo è obbligatorio *required* nel tipo base, non è possibile cambiarlo in *optional* nel tipo derivato (non è possibile avere un tipo derivato meno restrittivo del tipo base).

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation

Quando si deriva by extension da un tipo complesso con simple content, il tipo risultante sarà sempre un simple content.

L'unica cosa che è possibile fare è aggiungere attributi

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: simple content esempio

```
<xsd:complexType name='ExtendedPhoneType'>  
  <xsd:simpleContent> <xsd:extension  
    base='PhoneType'> <xsd:attribute  
      name="CallOnHolidays" type="xsd:string"/>  
    </xsd:extension> </xsd:simpleContent>  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation

Un tipo complesso *element-only content* può essere esteso (extended) o limitato (restricted) per generare un nuovo tipo.

Complex Type derivation

Derivare *by restriction* da un tipo *element-only content* vuol dire eliminare elementi o attributi.

Attenzione: gli elementi del tipo di base non vengono passati al tipo derivato (a differenza degli attributi)

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation

- Se gli elementi nel tipo di base sono dichiarati all'interno di una *sequence*, il tipo derivato non può cambiare tale dichiarazione in *all* oppure *choice*.
- Se il tipo di base ha specificato gli elementi con una dichiarazione *all*, il tipo derivato può cambiare la dichiarazione con *sequence*.
- Se il tipo di base ha specificato gli elementi con la dichiarazione *choice*, il tipo derivato deve avere ugualmente la dichiarazione *choice*.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation

- Se si vuole rimuovere un attributo bisogna dichiararlo nuovamente e specificarne l'uso *prohibited*.
- Un tipo derivato non può eliminare un attributo dichiarato obbligatorio nel tipo di base.
- Bisogna ridefinire un attributo nel tipo derivato, solo se si vuole limitare il suo spazio di valori (*gli attributi vengono ereditati dal tipo derivato*).

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:element name="Contact"  
type="RestrictedContactType"/>  
<xsd:complexType name='ContactType'  
<xsd:sequence> <xsd:element name="Phone"  
minOccurs="0"/> <xsd:element name="Email"  
minOccurs="0"/> </xsd:sequence> <xsd:attribute  
name="Name" /> <xsd:attribute name="Title"/>  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation: esempio

```
<xsd:complexType name='RestrictedContactType'>
  <xsd:complexContent>
    <xsd:restriction
      base='ContactType'>
      <xsd:sequence>
        <xsd:element
          name='Phone'>
          <xsd:simpleType>
            <xsd:restriction
              base='xsd:string'>
              <xsd:pattern
                value="[0-9]3-[0-9]3-[0-9]4"/>
            </xsd:restriction>
          </xsd:simpleType>
        </xsd:element>
      </xsd:sequence>
      <xsd:attribute name="Title" use="prohibited"/>
      <xsd:attribute name='Name' type='xsd:string'>
      </xsd:attribute>
    </xsd:restriction>
  </complexContent>
</complexType>
```


Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation

E' possibile derivare un *empty content* da un *element-only content*.

Ciò viene fatto mantenendo la *dichiarazione di restrizione vuota*.

Questo **non vale per gli attributi**, che si ereditano tutti.

Complex Type derivation

Gli elementi del tipo dell'elemento base devono essere opzionali

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type extension

*I tipi **complex type** vengono estesi per aggiungere elementi e attributi*

Complex Type extension

Quando si deriva per estensione, sia tutti gli elementi, sia tutti gli attributi si ereditano dal tipo base al tipo derivato.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:element name="Contact"  
type="ExtendedContactType"/> <xsd:complexType  
name="ContactType"> <xsd:sequence> <xsd:element  
name="Phone"/> <xsd:element name="Email"  
minOccurs="0"/> </xsd:sequence> <xsd:attribute  
name="Name" use="required"/> <xsd:attribute  
name="Title"/> </xsd:complexType> </xsd:schema>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:complexType name='ExtendedContactType'>  
  <xsd:complexContent>  
    <xsd:extension  
      base='ContactType'>  
      <xsd:sequence>  
        <xsd:element  
          name='Fax' />  
      </xsd:sequence>  
        <xsd:attribute  
          name='Department' />  
      </xsd:extension>  
    </xsd:complexContent>  
  </xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type: deriving

Si può derivare un *mixed content* sia by restriction sia by extension. Derivare un tipo complesso *mixed content* è simile a derivare un tipo complesso *element-only*.

Complex Type: deriving

deriving by restriction un mixed content type: *a mixed type*, *element-only* oppure *empty content*.

Gli elementi dichiarati nel tipo base non si ereditano, gli attributi si ereditano

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:complexType name="RestrictedNoteType"  
mixed='true'> <xsd:complexContent>  
<xsd:restriction base='NoteType'> <xsd:sequence>  
<xsd:element name="name"/> </xsd:sequence>  
</xsd:restriction> </xsd:complexContent>  
</xsd:complexType>
```

Rimuovendo l'attributo "mixed" dalla dichiarazione del tipo complesso si ottiene un element-only complex type da un tipo mixed content

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type: deriving - Mixed Content

Quando si deriva by restriction un mixed type, si può creare un *empty content type* se tutti gli elementi del tipo base sono optional.

Complex Type: deriving - Mixed Content

Per eliminare gli attributi c'è bisogno di dichiararli nel tipo derivato e di proibirne l'uso (*prohibited attribute*).

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation: esempio

```
<xsd:complexType name='SimpleNoteType'>  
  <xsd:simpleContent>  
    <xsd:restriction  
      base='NoteType'>  
      <xsd:simpleType>  
        <xsd:restriction base="xsd:string"/>  
      </xsd:simpleType>  
    </xsd:restriction>  
  </xsd:simpleContent>  
</xsd:complexType>
```

Le specifiche XSD permettono di derivare un simple content da un mixed-content type

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:complexType name='NoteType'  
mixed='true'> <xsd:sequence> <xsd:element  
name="name"/> <xsd:element name="mobile"/>  
</xsd:sequence> </xsd:complexType>
```

*Quando si deriva by **extension**, possiamo solo generare **mixed content complex type***

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation: esempio

```
<xsd:element name="InvoiceNote"  
type="ExtendedNoteType"/>
```

```
  <xsd:complexType name='ExtendedNoteType'  
mixed='true'> <xsd:complexContent>  
  <xsd:extension base='NoteType'> <xsd:sequence>  
    <xsd:element name='email'> </xsd:sequence>  
  </xsd:extension> </xsd:complexContent>  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<InvoiceNote> Chiamare <name>Angelo</name> sul  
numero di cellulare <mobile>0039 321  
32321</mobile> or  
<email>angelo.delgrosso@ilc.cnr.it</email> se  
l'ordine non va a buon fine. </InvoiceNote>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type extension

**si può derivare un nuovo tipo da un elemento di tipo
empty content sia by restriction sia by extension**

Complex Type extension

Un elemento con *empty content* ha solo attributi, l'unico motivo per derivare da questa tipologia di content model by restriction è quella di eliminare attributi oppure aggiungere restrizioni.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type extension

Quando si deriva by extension è possibile generare empty type, mixed type oppure element only content type.

Elementi per la definizione degli schemi xml

principi XSD

Complex Type derivation: esempio

```
<xsd:complexType name='ElemOnlyPhoneType'>  
  <xsd:complexContent>  
    <xsd:extension  
      base='EmptyPhoneType'>  
      <xsd:sequence>  
        <xsd:element name='Mobile' />  
      </xsd:sequence>  
    </xsd:extension>  
  </xsd:complexContent>  
</xsd:complexType>
```

Si può derivare un element only content type da un empty content type by extension

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:complexType name='MixedPhoneType'  
mixed='true'> <xsd:complexContent>  
  <xsd:extension base='EmptyPhoneType'>  
    <xsd:sequence> <xsd:element name="Mobile"/>  
  </xsd:sequence> </xsd:extension>  
</xsd:complexContent> </xsd:complexType>
```

Si può derivare un mixed content type da un empty content type by extension

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<Phone Home='‘888-888-8888’’,  
Work='‘777-777-7777’’> Se non disponibile  
contattare il cellulare al numero  
<Mobile>666-666-6666</Mobile>. </Phone>
```


Elementi per la definizione degli schemi xml

principi XSD

Derivazione: riepilogando

Base Content Type	Derived Content Type							
	Simple		Element Only		Mixed		Empty	
	Rest	Ext	Rest	Ext	Rest	Ext	Rest	Ext
Simple	Yes	Yes						
Element Only			Yes	Yes			Yes	
Mixed			Yes		Yes	Yes	Yes	
Empty				Yes		Yes	Yes	Yes

Nuovi complex types possono essere creati a partire da altri complex types by extension and restriction.

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: final attribute

La derivazione di un tipo complesso può essere controllata impiegando l'attributo final nella dichiarazione del tipo

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: final attribute

L'attributo final può assumere i seguenti valori:

- **restriction**: limita la derivazione by restriction
- **extension**: limita la derivazione by extension
- **#all**: limita la derivazione by restriction e by extension

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio

```
<xsd:complexType name='ContactType'  
  final='extension restriction'> <xsd:sequence>  
  <xsd:element name="Phone" minOccurs="0"/>  
  </xsd:sequence> <xsd:attribute name="Name" />  
</xsd:complexType>
```

Elementi per la definizione degli schemi xml

principi XSD

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Complex Type derivation: esempio - derivazione non consentita

```
<xsd:complexType name='ExtendedContactType'>  
  <xsd:complexContent>  
    <xsd:extension  
      base="ContactType"/>  
  </xsd:complexContent>  
</xsd:complexType>
```

Progress status

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

1 XML Schema Definition (XSD)

2 Introduzione Text Encoding Initiative

3 Conclusioni

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Motto

TEI: Yesterday's information tomorrow

Dal sito TEI

“an international and interdisciplinary standard that enables libraries, museums, publishers, and individual scholars to represent a variety of literary and linguistic texts for online research, teaching, and preservation”

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Testo di riferimento

Guidelines for Electronic Text Encoding and Interchange
(<http://www.tei-c.org/Guidelines/>)

testo di ausilio

BURNARD, Lou. What is the Text Encoding Initiative? How to add intelligent markup to digital resources. Nouva edizione [online]. Marseille: OpenEdition Press, 2014 (creato il 13 ottobre 2018). Disponibile su Internet:
<http://books.openedition.org/oep/426>. ISBN:
9782821834606. DOI: 10.4000/books.oep.426.

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

un po' di storia

- 1987: necessità di standard che permetta la creazione e l'interscambio di documenti per mezzo di archivi informatici(convegno NY)
- 1990: prima versione delle Guidelines (TEI P1)
- 1990-94: fondi garantiti da enti quali NEH, Mellon Foundation, la Comunità Europea; supporto di ACH, ACL, ALLC

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

un po' di storia

- 2000: nascita del TEI Consortium, associazione non profit per lo sviluppo dello standard TEI
- 2002: passaggio da SGML a XML con la v. P4
- 2007: nuova versione TEI P5, continuamente aggiornata

TEI Guidelines

versioni P1 e P3 basate su SGML
versione P4

- standard precedente, ancora impiegata
- basata su XML, DTD tradizionale
- pubblicata in forma definitiva nel 2002
- <http://www.tei-c.org/Guidelines/P4/>

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Guidelines: versione P5

- basata su XML, schema RelaxNG (e DTD tradizionale)
- pubblicata alla fine del 2007, aggiornata due volte l'anno
- molte novità interessanti (in particolare: maggior modularità)
- <http://www.tei-c.org/Guidelines/P5/>

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Guidelines: Obiettivi

- better interchange and integration of scholarly data
- support for all texts, in all languages, from all periods
- guidance for the perplexed: what to encode - hence, a user-driven codification of existing best practice
- assistance for the specialist: how to encode — hence, a loose framework into which unpredictable extensions can be fitted

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Guidelines: Obiettivi

These apparently incompatible goals result in a highly flexible, modular, environment for DTD customization. Lou Burnard, TEI and XML: a marriage made in heaven? (<http://www.tei-c.org/Talks/marriage.xml?style=printable>)

Che cosa offre la TEI

- un ricco (e complesso) manuale di codifica, (le Guidelines for Electronic Text Encoding and Interchange)
- un numero elevato di elementi (sia strutturale sia semantico)
- schemi di codifica
- infrastruttura modulare e personalizzabile

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Che cosa offre la TEI

- **è possibile scegliere soltanto i moduli necessari**
- **è possibile modificare le definizioni degli elementi**

Intro Text Encoding Initiative

TEI

Supporto per gli utenti

- il sito del consorzio (<http://www.tei-c.org/>)
- pagine relative alle varie versioni delle Guidelines
- software, tutorial, ecc.
- il wiki:
http://www.tei-c.org/wiki/index.php/Main_Page
- la mailing list TEIL
- Github: <https://github.com/TEIC>
- TEI by example: <http://teibyexample.org/>

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Novità della versione P5

- modulo di descrizione dei manoscritti
- grafica e multimedia
- standoff markup
- supporto per i namespace XML
- miglioramenti nel modulo feature structure

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Novità della versione P5 (cont.)

- modulo di descrizione dei manoscritti
- grafica e multimedia
- standoff markup
- supporto per i namespace XML
- miglioramenti nel modulo feature structure

TEI: Struttura modulare

- si scelgono soltanto i moduli che corrispondono alle proprie esigenze, in modo da realizzare rapidamente uno schema di codifica appropriato
- ogni modulo contiene un certo numero di elementi (tagset)
- gli elementi sono organizzati in classi (strutturali, semantiche)
- gli attributi sono organizzati in classi (globali e specifici)

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI: Struttura modulare - Moduli essenziali

- **tei**: definisce le classi di elementi, le macro e i datatype che verranno usati per tutti i moduli
- **header**: l'intestazione contenente i metadati relativi al documento TEI XML
- **textstructure**: elementi strutturali per qualsiasi tipo di testo
- **core**: elementi utili in qualsiasi tipo di documento

TEI: Struttura modulare - Moduli facoltativi

- **analysis**: strumenti per analisi (linguistica etc.) del testo
- **corpus**: gestione di corpora linguistici
- **drama**: elementi per testi teatrali e drammatici
- **gaiji**: rappresentazione di caratteri e glifi non standard
- **msdescription**: metadati relativi a manoscritti

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI: Struttura modulare - Moduli facoltativi

- **spoken**: trascrizione del parlato
- **textcrit**: apparato critico
- **transcr**: trascrizione di fonti primarie (manoscritti)
- **verse**: elementi supplementari per testi poetici

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Lite

una specifica **personalizzazione** della TEI versione P4/5

*A simple demonstration of how the TEI encoding scheme might be adopted to meet 90% of the needs of 90% of the TEI user community (dalla Prefatory Note:
[urlhttp://www.tei-c.org/Lite/](http://www.tei-c.org/Lite/))*

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Lite

La **versione P4** è stata tradotta anche in italiano: TEI Lite: introduzione alla codifica dei testi, a cura di *F. Ciotti* (http://www.tei-c.org/Lite/teiu5_it.html).

Molto usata, ma presenta varie limitazioni, con la versione P5 è più semplice produrre una versione semplificata o personalizzata

TEI Pizza chef

la TEI P4 poteva essere modificata usando un programma su web chiamato Pizza chef

<http://www.tei-c.org.uk/pizza.html>.

Metafora della base e dei condimenti (toppings), corrispondenti ai moduli indispensabili e a quelli facoltativi.

Meccanismo efficace, soprattutto considerando l'alternativa (modifica manuale delle DTD TEI).

Tool obsoleto, necessaria comunque modifica manuale per nuovi elementi

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Roma

Per la P5 si è deciso di proporre una modularizzazione più efficace, e uno strumento di personalizzazione più potente. Strumento basato sul nuovo formato ODD.

TEI Roma

Il metodo da seguire per la versione P5 (quella che useremo) fino all'arrivo del successore (Byzantium)

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Roma

- possibilità di scegliere, escludere, modificare sia gli elementi (e le classi di elementi), sia gli attributi (e le classi di attributi)
- possibilità di aggiungere elementi (eventualmente inserendoli nelle classi preesistenti)
- possibilità di salvare lo schema in tre formati diversi: DTD tradizionale, W3C e RelaxNG (anche in forma compatta)

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Roma: Il formato ODD

le versioni P1 – P4 delle DTD TEI erano nel formato DTD language e dipendevano dalla sintassi SGML per molti aspetti

TEI Roma: Il formato ODD

One Document Does it all (ODD): set di specifiche in base alle quali un semplice documento TEI XML *“defines a schema in terms of the modules it requires, together with any possible modifications, such as the desired root element”*.

Risultato finale: lo schema nel linguaggio desiderato e la relativa documentazione

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI Roma: Il formato ODD

tutorial: Getting Started with P5 ODDs

http:

[//www.tei-c.org/Guidelines/Customization/odds.xml](http://www.tei-c.org/Guidelines/Customization/odds.xml)

Il futuro della TEI

negli ultimi anni gli schemi TEI sono stati oggetto di alcune critiche

Il futuro della TEI

- la TEI è troppo grande / complicata / piccola
- la TEI è basata su XML e questo formato è in declino
- la TEI/XML non supporta le gerarchie multiple
- la TEI non supporta il markup di tipo stand-off

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Il futuro della TEI

la cosa importante da ricordare è che il formato XML non è la TEI (in passato SGML, in futuro chissà → abstraction level) al contrario, se c'è una discrepanza fra Guidelines e schemi, la precedenza va alle Guidelines

Intro Text Encoding Initiative

TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

text encoding con la TEI

E' caldamente raccomandato usare direttamente la versione più recente della P5.

La flessibilità della P5 permette di definire uno schema di codifica che corrisponda precisamente al modello

La comunità di utenti e sviluppatori TEI offre un buon supporto.

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Un documento TEI P5 'minimo'

- prologo XML
- intestazione TEI
- elementi strutturali
- elementi semantici dei moduli base

Moduli di base: tei, header, textstructure, core

un documento TEI P5 'minimo'

Anche usando soltanto i moduli essenziali si ha a disposizione uno schema adatto alla marcatura di numerosi tipi di testi.

Schemi "leggeri" consigliati: la TEI Lite, o se necessario una versione più ridotta della P5 (TEI Absolutely Bare)

Caratteristiche degli elementi illustrati

- gli elementi TEI rientrano nelle categorie generali di elementi XML che abbiamo visto
- elementi che possono contenere solo altri elementi (= elementi strutturali)
- elementi che possono contenere altri elementi e testo
- elementi che possono contenere solo testo
- elementi vuoti (es. `<pb/>`)
- gli elementi vuoti marcano una gerarchia differente

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Gerarchie multiple

```
<?xml version="1.0" encoding="utf-8"?> <text>
<titolo>Gli assassini della Rue Morgue</titolo>
<intestazione> I </intestazione> <pagina n='5'>
<p>Le facoltà mentali che si sogliono chiamare
analitiche sono, di per se stesse, poco
suscettibili di analisi [...]</p> <p>La facoltà di
risolvere è probabilmente molto rinfor- </pagina>
<pagina n='6'> zata dallo studio delle
matematiche e in modo particolare dell'altissimo
ramo di questa scienza che[...] </p> </pagina>
</text>
```

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Gerarchie multiple- elementi vuoti

```
<?xml version="1.0" encoding="utf-8"?> <text>  
<titolo>Gli assassini della Rue Morgue</titolo>  
<intestazione> I </intestazione> <pagina n="5"/>  
<p>Le facoltà mentali che si sogliono chiamare  
analitiche sono, di per se stesse [...]</p> <p>La  
facoltà di risolvere è probabilmente molto rinfor-  
<pagina n="6"/> zata dallo studio delle  
matematiche e in modo particolare dell'altissimo  
ramo di questa scienza che [...] </p> </text>
```

Struttura di un documento TEI

- *struttura fondamentale all'interno della radice (<TEI>)*
- una intestazione TEI (<teiHeader>)
- un testo: <text> (o più testi, cfr. infra)

Contenuto del TEI header

- metadati relativi al documento (utili per collezioni di testi codificati)
- descrizione del file usando <fileDesc> (obbligatoria)
- descrizioni relative al tipo di codifica, al contenuto del documento, alle sue revisioni (facoltative)

E' possibile includere testi introduttivi e spiegazioni relative alla codifica effettuata (preziosi per l'interscambio!)

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Esempio

```
<?xml version="1.0" encoding="utf8"?> <!DOCTYPE  
TEI SYSTEM ‘‘tei_lite.dtd’’> <TEI  
xmlns=‘‘http://www.teic.org/ns/1.0’’> <teiHeader>  
</teiHeader> <text> <div> <p></p> </div> </text>  
</TEI>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Esempio

```
<?xml version="1.0" encoding="utf-8"?> <?xml-model  
href="tei-lite.rng"?> <TEI  
xmlns='http://www.tei-c.org/ns/1.0''>  
<teiHeader>...</teiHeader> <text> <div> <p></p>  
</div> </text> </TEI>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

documento TEI - schema di intestazione TEI minima

```
<teiHeader> <fileDesc> <titleStmt>...</titleStmt>  
<publicationStmt>...</publicationStmt>  
<sourceDesc>...</sourceDesc> </fileDesc>  
</teiHeader>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

documento TEI - schema di intestazione TEI minima

Metadati essenziali riguardo il titolo, la modalità di diffusione e la fonte originaria di un testo codificato.

Permettono classificazione, archiviazione ed elaborazione bibliografica

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

```
<teiHeader> <fileDesc> <titleStmt> <title>La  
Divina Commedia:  versione elettronica</title>  
<respStmt> <resp>Conversione TEI P5 a cura  
di</resp><name>M. Rossi</name> </respStmt>  
</titleStmt> <publicationStmt>  
<publisher>Università di Pisa</publisher>  
<date>2002-11-07</date> <availability  
status='‘restricted’'><p></p></availability>  
</publicationStmt> <sourceDesc> <bibl><title>La  
Divina Commedia</title><author>Dante Alighieri  
</author><publisher>Mondadori</publisher>  
<date>1988</date></bibl> </sourceDesc> </fileDesc>  
</teiHeader>
```

Le altre componenti dell'intestazione TEI

- `<encodingDesc>` informazioni riguardo lo schema (e il modello di codifica) utilizzato
- `<profileDesc>` descrizione del testo: quando è stato creato, da chi, usando quali lingue etc.
- `<revisionDesc>` informazioni sulle versioni del file

I metadati sono una componente essenziale di qualsiasi progetto di digitalizzazione

Elementi strutturali

- `<text>` un singolo testo di qualsiasi tipo (punto di partenza della gerarchia).
- `<facsimile>` riproduzione della fonte primaria, può affiancare o sostituire `<text>`
- `<front>` figlio di `<text>` materiale che precede il testo
- `<body>` figlio di `<text>` rappresenta il testo stesso
- `<back>` figlio di `<text>` materiale che segue il testo
- `<group>` figlio di `<text>` alternativo a `<body>`, raggruppa testi diversi

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Esempio schematico di documento TEI

```
<TEI> <teiHeader> [informazioni del TEI Header]
</teiHeader> <text> <front> [premessa, dedica ...]
</front> <body> [corpo del testo ...] </body>
<back> [postfazione, appendice ...]</back> </text>
</TEI>
```


Costruire documenti compositi

- rimpiazzando il `<body>` con un gruppo (`<group>`) di testi si ottiene un documento composito
- ciascuno di questi testi è rappresentato secondo una struttura standard
- un'altra possibilità è creare un corpus con `<teiCorpus>`
- intestazioni (`<teiHeader>`) separate per il corpus e per ciascun gruppo di testi
- struttura più complessa, su più livelli

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

```
<TEI> <teiHeader> [ intestazione del testo  
composito ] </teiHeader> <text> <front> [ front  
matter del composito ] </front> <group> <text>  
<front> [ front matter del primo testo ] </front>  
<body> [ body del primo testo ] </body> <back> [  
back matter del primo testo ] </back> </text>  
<text> <front> [ front matter del secondo testo]  
</front> <body> [ body del secondo testo ] </body>  
<back> [ back matter del secondo testo ] </back>  
</text> ... [ altri testi o gruppi di testi ] ...  
</group> <back> [ back matter del composito ]  
</back> </text> </TEI>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Costruzione di corpora TEI

```
<teiCorpus> <teiHeader> [metadati per il corpus]
</teiHeader> <TEI> <teiHeader> [metadati relativi
al I testo]</teiHeader> <text> [primo testo del
corpus] </text> </TEI> <TEI> <teiHeader>[metadati
relativi al II testo]</teiHeader> <text> [secondo
testo del corpus] </text> </TEI> </teiCorpus>
```

Altri elementi strutturali fondamentali

- suddivisioni del testo, non numerati: `<div>` (nessun limite di nidificazione)
- suddivisioni del testo, numerati: `<div1> ... <div7>` (massimo 7 livelli)
- paragrafi: `<p>`
- testo riferito: `<q>` (discorso diretto, citazioni, etc.)

Altri elementi strutturali fondamentali

- versi: strofe `<lg>` e singoli versi `<l>`
- testi teatrali: discorsi `<sp>` che possono contenere paragrafi `<p>` o versi `<l>`, oltre a direzioni di scena `<stage>`
- milestone tags: `<pb/>`, `<lb/>`, `<cb/>`, `<milestone/>`
- notare che un `<div>` può contenere un `<floatingText>` (possibilità di introdurre gerarchie complesse).

Apertura e chiusura di un <div>

- <head>: qualunque tipo di intestazione: il titolo di un'opera, l'intestazione di un paragrafo, di una sezione, ecc.
- l'attributo *type* permette di classificare in base a una tipologia
- <epigraph> citazione all'inizio del testo, o nella pagina del titolo, eventualmente con riferimento bibliografico
- <opener> raggruppa una serie di elementi (data, luogo, saluti, ecc.) all'inizio del <div>, specie di una lettera

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Apertura e chiusura di un <div>

- <argument>: lista degli argomenti trattati nel <div>
- <trailer> frase che compare alla fine del <div> (ad esempio “Fine del capitolo 1”)
- <closer> raggruppa un serie di elementi (data, luogo, saluti, etc.) alla fine del <div>, specie di una lettera

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

```
<div type="lettera"> <opener> <dateline> <name  
type='‘place’>Pisa</name> <date>20 marzo  
2015</date> </dateline> <salute>Gentilissima  
Prof.ssa Scannagatti,</salute> </opener> <p>sono  
spiacente di doverle comunicare che un'invasione  
di cavallette si è abbattuta sui miei quaderni  
incautamente lasciati in giardino, e li ha  
divorati interamente.</p> <p>Questo purtroppo  
significa che non posso mostrare i compiti svolti,  
come sempre, con solerzia e assiduo impegno.</p>  
<closer> <salute>Certo di poter contare sulla sua  
comprensione le porgo i miei migliori  
saluti,</salute> <signed>Pierino Rossi</signed>  
</closer> </div>
```


Errori frequenti

Si fraintende il significato dell'elemento <fileDesc>

- serve in primo luogo a dare informazioni sul file stesso, non sul testo originale
- il riferimento alla fonte dalla quale è tratto il testo codificato deve essere inserito nel <sourceDesc>

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Errori frequenti

I titoli sono codificati con `<title>` soltanto nel caso di riferimenti bibliografici i titoli del testo, dei capitoli etc. si marcano con `<head>`

Nota sugli errori possibili

Tre categorie:

- **errori sintattici:** un elemento inserito in un punto sbagliato della gerarchia, o che non può contenere testo etc.
- **errori di marcatura semantica:** usare un elemento inadatto allo scopo, ad esempio marcare un titolo con `<jemph>`
- **errori di interpretazione** del testo (che portano al II tipo o all'assenza del markup che andrebbe inserito)

Gli errori del primo tipo sono i più facili da individuare e

Alcuni attributi possono essere usati con qualsiasi elemento (v. la classe `att.global`)

Attributi globali

- **n** un numero o un nome non univoco, possibilmente breve, per identificare un elemento
- **rend** informazioni relative all'aspetto (*originale!*) del testo
- **rendition** simile a *@rend*, ma fa riferimento a elementi `<rendition>` inseriti nell'`<encodingDesc>` (dentro `<tagsDecl>`)

Attributi globali

- **xml:lang** la lingua del testo contenuto da un elemento
- **xml:id** un identificatore univoco per l'elemento

NOTA: in base ai moduli usati nello schema sono disponibili ulteriori attributi globali

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Esempio

```
<text> <body> <div n="ch1" type='‘chapter’> <pb  
n="1"/>[...] <p>[...] risulta chiaro se avete  
letto <title rend="underline" xml:lang='‘fra’>Les  
fleurs du mal</title> [...]</p> <p>[...] un  
grande esempio di <foreign xml:lang='‘fra’>savoir  
faire</foreign> [...]</p> [...] </div> [ altri  
div ... ] </body> </text>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Esempio

```
<text> <body> <div n="ch1" type="chapter"> <pb  
n="1"/> [...] <p n="1">[...] descritto altrove  
(si veda ad esempio <ref target="#Rossi94">Rossi  
1994</ref>) [...] </p> [...] </div> [ altri div  
... ] <div n="bib" type="bibliography"> [...] <bibl  
xml:id="#Rossi94"> <author>Rossi,  
M.</author>[...]</bibl> [...] </div> </body>  
</text>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Errori frequenti

`<div>` non può essere usato allo stesso livello gerarchico di `<p>`, in altre parole non si può alternare `<div>` con `<p>`

Errore!

```
<div> [...] </div> <p> [...] </p> <div> [...]
</div>
```


Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Errori frequenti

`<div>` e tutti gli altri elementi strutturali *puri* non possono contenere testo.

Errore!

```
<div>Pippo</div> <person>Pippo</person>
```

Enfasi e termini particolari

- `<emph>` parole o frasi enfatizzate nel testo. (Questo è il `<emph>mio</emph>` computer!)
- `<foreign>` parola o frase in una lingua diversa. (In quel punto entrò il bidello a dare il `<foreign xml:lang='lat'>finis</foreign>`).

Enfasi e termini particolari

- `<distinct>` “diverso” dal testo perché arcaico, gergale, ecc. (Saltò in groppa al `<distinct>`fido destriero`</distinct>`)
- `<hi>` elemento generico. (`<hi rend=“double”>`N`</hi>`el mezzo del cammin di nostra vita. Il suo nome è `<hi rend=“italic”>`Mario Rossi`</hi>`)

Enfasi e termini particolari

- `<mentioned>` parola o frase menzionata ma non usata.
(Il termine corretto è
`<mentioned>epigrafe</mentioned>`)
- `<soCalled>` parola o espressione da cui ci si distanzia (il
cosiddetto `<soCalled>darwinismo
sociale</soCalled>`)

Enfasi e termini particolari

- `<term>` una o più parole considerate termine tecnico.
(Possiamo definire il `<term xmlid="NPL" rend="‘‘italic’’>neopositivismo logico</term>`)
- `<gloss>` una spiegazione o glossa riguardo il testo.
(`<gloss target=‘‘#NPL’’>una corrente filosofica basata sul principio che la filosofia debba aspirare al rigore proprio della scienza</gloss>`)

Esercizio

Marcare un testo plain text di circa 3000 caratteri a piacere.

- inserire prologo XML
- marcare la struttura usando gli elementi fin qui descritti in particolare marcare tutti i paragrafi usando `<p>` e la struttura editoriale usando `¶div`
- verificare che sia ben formato con `xmllint`
- salvare il file XML su github

Progress status

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

1 XML Schema Definition (XSD)

2 Introduzione Text Encoding Initiative

3 Conclusioni

Intro Text Encoding Initiative TEI

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

Nota sugli elementi illustrati

- punto di riferimento per il corso: la TEI P5 “completa”
- iniziamo con la TEI Lite (= versione “leggera”)
- proposti elementi di uso generale e di uso specifico
- per cominciare, la TEI Lite è più semplice da gestire
- elementi non presenti di default evidenziati

Nota sugli elementi illustrati

- in caso di dubbio consultare le liste degli elementi:
- TEI Lite: <http://www.tei-c.org/release/doc/tei-p5-exemplars/html/teilight.doc.html>
- TEI P5: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/REF-ELEMENTS.html>

NOTA: *quando si crea lo schema controllare che gli elementi siano effettivamente inclusi nei moduli prescelti*

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Codifica di
Testi - XML
Schema e
Intro TEI
a.a.
2018-2019

A.M. Del
Grosso

XML Schema
Definition
(XSD)

Introduzione
Text Encoding
Initiative

Conclusioni

TEI

Suggerimenti riguardo gli esercizi non fate copia e incolla da Wordpad o Notepad, aprite direttamente i testi usando l'editor i documenti XML hanno come estensione .xml, quindi salvate subito una copia con tale estensione aprite le Guidelines TEI nel browser per prima cosa marcare la struttura: cominciare la marcatura semantica come secondo passo ogni esercizio ha degli obiettivi da raggiungere: controllate prima di spedire non lasciate accumulare gli errori