

# Dalla Recensio all'Emendatio Digitale

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

**Teoria, Prassi e Strumenti**

Istituto di Linguistica Computazionale “A. Zampolli”,  
15th June 2020

# Argomenti trattati

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

5 Strumenti per edizioni critiche

6 Considerazioni Finali

7 References

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

**1** Presentazione

**2** Panoramica Text Encoding Initiative

**3** Panoramica eXtensible Markup Language

**4** TEI: Codifica Apparato Critico

**5** Strumenti per edizioni critiche

**6** Considerazioni Finali

# Di cosa mi occupo

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Filologia Digitale e Computazionale

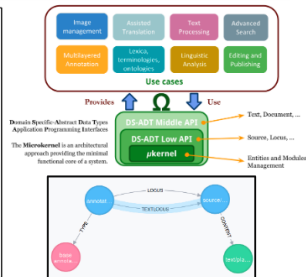
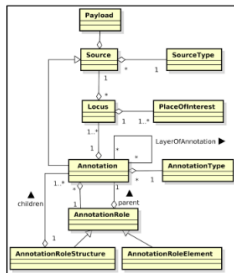
Attività di ricerca per lo sviluppo di sistemi di linguistica e filologia digitale e computazionale volti alla produzione, rappresentazione, analisi, fruizione e interrogazione di testi di tradizione medievale, a stampa e di autori moderni e contemporanei.

# Profilo professionale e di ricerca

## In sintesi

Ingegnere Informatico prestato alla filologia computazionale

```
text = Text.of("Literary Text to process", URI.create("//source/text/000"));  
annotation = AnnotationText.of("Annotation on the text",  
    URI.create("//annotation/text/123"));  
annotation.addLocus(text, 13, 18);  
annotation.save();
```



# Temi del seminario

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## ■ Filologia del testo

- disciplina storica (insieme di discipline) dedicata al recupero/ricostruzione di un'opera - per lo più letteraria - e allo studio critico delle sue testimonianze condotto con metodo scientifico.

## ■ Metodo di Lachman

## ■ Apparato Critico

- Apparato critico: nella edizione critica di un testo, l'apparato critico è il luogo (che può essere: a piè di pagina, in appendice al testo oppure dopo la nota al testo) in cui l'editore accoglie - a volte discutendolo (e comunque offrendo la possibilità di verifica del suo lavoro critico) - il complesso delle annotazioni, correzioni e varianti portate dalla tradizione e da lui giudicate erronee (o non meritevoli di essere accolte come lezione a testo).

# Temi del seminario

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## ■ Codifica del testo

- Rappresentazione formale con tecnologia digitale
- Livello di rappresentazione dei caratteri
- Livello di rappresentazione della struttura e dei fenomeni testuali

# Temi del seminario

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

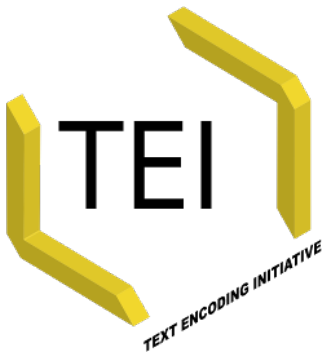
Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References





# Obiettivo del seminario

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

approfondiremo

Tecniche per la rappresentazione digitale di edizioni critiche adottando le specifiche e le linee guida della *Text Encoding Initiative* implementate adottando l'eXtensible Markup Language (*TEI-XML*).

approfondiremo

Strumenti per la collazione automatica e la pubblicazione di edizioni critiche digitali

# Perché è importante la codifica dei testi

## Motivazioni pratiche

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Perché codificare i testi

Per rendere disponibile l'immenso patrimonio testuale tramite l'uso di sistemi digitali e computazionali è necessario effettuare una trasposizione/transcodifica\* dei testi dal loro supporto originario verso il nuovo supporto elettronico (*Machine Readable Form and Machine Actionable Form*).

\* *procedimento di conversione dei dati codificati secondo un sistema verso un sistema diverso*

# Elementi di Codifica dei Caratteri

## Definizioni

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### Rappresentare il testo in formato digitale

L'adozione di metodologie e tecnologie informatiche per il trattamento di documenti testuali richiede in primo luogo la disponibilità di un'adeguato sistema di rappresentazione digitale dei dati - presenti nella risorsa originale - e conseguentemente un formalismo adatto a tale rappresentazione.

# Elementi di Codifica del testo

## Formalismi

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### Formati e formalismi di codifica

Ogni pezzo di informazione aggiunta ad un testo grezzo attraverso l'inserimento di dati metatestuali (markup, annotazione, codifica), costituisce il risultato di una analisi e di una interpretazione che è stata condotta (da un umano o da una macchina) al fine di esplicitare e rappresentare nel modo più accurato e completo possibile le informazioni da veicolare attraverso il formato digitale prescelto (anche in modo incrementale).

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

5 Strumenti per edizioni critiche

6 Considerazioni Finali

# Markup language e XML

soluzione corrente per la codifica dei testi

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## TEI-XML

Considerato ad oggi lo standard de facto per la codifica dei testi è lo schema XML messo a punto dalla Text Encoding Initiative (TEI-XML).

### Perché TEI

La Text Encoding Initiative (TEI) è un autorevole progetto internazionale, a cui afferiscono varie organizzazioni e università, il cui scopo è fornire agli studiosi di informatica umanistica uno strumento il più espressivo e flessibile possibile per rappresentare qualsiasi aspetto di interesse relativo alla risorsa testuale da rappresentare digitalmente.

## Qual è l'obiettivo della TEI

L'obiettivo della TEI è quello di fornire linee guida per la creazione e la gestione in forma digitale di qualsiasi tipo di dato creato e usato in ambito umanistico.

E per questo motivo il consorzio investe molte risorse per la accessibilità e la divulgazione della tecnologia che da anni sviluppa.



# I principi fondamentali della TEI



- Le linee guida della TEI privilegiano il “significato” (meaning) del testo piuttosto che l’“aspetto” (layout); privilegia il modello del testo, piuttosto che il formato.
- La TEI è stata progettata per essere indipendente dagli strumenti software che la usano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche dalla propria comunità di riferimento (community-driven).

# I principi fondamentali della TEI



- Le linee guida della TEI privilegiano il “significato” (meaning) del testo piuttosto che l’“aspetto” (layout); privilegia il modello del testo, piuttosto che il formato.
- La TEI è stata progettata per essere indipendente dagli strumenti software che la usano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche dalla propria comunità di riferimento (community-driven).

# I principi fondamentali della TEI



- Le linee guida della TEI privilegiano il “significato” (meaning) del testo piuttosto che l’“aspetto” (layout); privilegia il modello del testo, piuttosto che il formato.
- La TEI è stata progettata per essere indipendente dagli strumenti software che la usano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche dalla propria comunità di riferimento (community-driven).

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### 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 octobre 2018). Disponibile su Internet:  
<http://books.openedition.org/oep/426>. ISBN:  
9782821834606. DOI: 10.4000/books.oep.426.

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### 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

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>)

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Che cosa offre la TEI

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

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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](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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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](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*

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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)

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

### Un documento TEI P5 'minimo'

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

# Intro Text Encoding Initiative

## TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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)*

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References



# Intro Text Encoding Initiative

Gerarchie multiple

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<?xml version="1.0" encoding="utf-8"?>
<text>
  <titolo>Gli assassini della Rue Morgue</titolo>
  <intestazione> I </intestazione>

  <pagina n="5">
    <p>Le facolt'a mentali che si sogliono chiamare analitiche sono, di per se
    stesse, poco suscettibili di analisi [...]</p>
    <p>La facolta` di risolvere `e 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>
```

# Intro Text Encoding Initiative

## Schemi di codifica TEI – Gerarchie multiple

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<?xml version="1.0" encoding="utf-8"?>
<text>
  <titolo>Gli assassini della Rue Morgue</titolo>
  <intestazione> I </intestazione>

  <pagina n="5"/>
  <p>
    Le facolt`a mentali che si sogliono chiamare analitiche sono, di per se
    stesse [...]
  </p>
  <p>
    La facolt`a di risolvere `e probabilmente molto rinfor-
    zata dallo studio delle matematiche e in modo particolare dell'altissimo
    ramo di questa scienza che [...]
  </p>
</text>
```

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Struttura di un documento TEI

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

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<?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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

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

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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 – Intestazione TEI minima

```
<teiHeader>
...
  <fileDesc>
    <titleStmt>
      <!--
      |   Sezione per la codifica del titolo
      |   dell'edizione digitale
      -->
    </titleStmt>
    <publicationStmt>
      <!--
      |   Sezione per la codifica dei
      |   dati di diffusione e distribuzione dell'edizione
      -->
    </publicationStmt>
    <sourceDesc>
      <!--
      |   Descrizione della fonte primaria
      -->
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References



# Intro Text Encoding Initiative

## Schemi di codifica TEI – Moduli base

```
<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>
```

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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*

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

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

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<TEI>
...
  <teiHeader> <!-- intestazione --> </teiHeader>
  <text>
    <front> <!-- frontespizio testo composito --> </front>
    <group>
      <text>
        <front> <!-- frontespizio del primo testo --> </front>
        <body> <!-- body del primo testo --> </body>
        <back> <!-- appendice del primo testo --> </back>
      </text>
      <text>
        <front> <!-- frontespizio del secondo testo --> </front>
        <body> <!-- body del secondo testo --> </body>
        <back> <!-- appendice del secondo testo --> </back>
      </text> <!-- altri testi o gruppi di testi -->
    </group>
    <back> <!-- appendice del testo composito -->
  </back>
</text>
</TEI>
```

# Intro Text Encoding Initiative

## Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<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>
```

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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.)



# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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).

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<text><body>
...
  <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>
</body>
</text>
```

### 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

## 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*

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`)



# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

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

# Intro Text Encoding Initiative

## Schemi di codifica TEI: Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

```
<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>
    <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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Errori frequenti

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

## Errore!

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

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`).

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`)

# Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`)



# Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## 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>`)

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

## 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

## TEI framework

La tecnologia TEI ha un framework concettuale diviso in

- Moduli
- Classi
- Macro
- Tipi di Dato

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Moduli TEI

Un modulo è semplicemente un contenitore per una serie di dichiarazioni uniformi e coerenti per gli elementi TEI e le relative classi.

# Infrastruttura TEI

## Tabella Moduli TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

Module name	Formal public identifier	Where defined
analysis	Analysis and Interpretation	17. <i>Simple Analytic Mechanisms</i>
certainty	Certainty and Uncertainty	21. <i>Certainty, Precision, and Responsibility</i>
core	Common Core	3. <i>Elements Available in All TEI Documents</i>
corpus	Metadata for Language Corpora	15. <i>Language Corpora</i>
dictionaries	Print Dictionaries	9. <i>Dictionaries</i>
drama	Performance Texts	7. <i>Performance Texts</i>
figures	Tables, Formulae, Figures	14. <i>Tables, Formulæ, Graphics and Notated Music</i>
gaiji	Character and Glyph Documentation	5. <i>Characters, Glyphs, and Writing Modes</i>
header	Common Metadata	2. <i>The TEI Header</i>
iso-fs	Feature Structures	18. <i>Feature Structures</i>
linking	Linking, Segmentation, and Alignment	16. <i>Linking, Segmentation, and Alignment</i>
msdescription	Manuscript Description	10. <i>Manuscript Description</i>
namesdates	Names, Dates, People, and Places	13. <i>Names, Dates, People, and Places</i>
nets	Graphs, Networks, and Trees	19. <i>Graphs, Networks, and Trees</i>
spoken	Transcribed Speech	8. <i>Transcriptions of Speech</i>
tagdocs	Documentation Elements	22. <i>Documentation Elements</i>
tei	TEI Infrastructure	1. <i>The TEI Infrastructure</i>
textcrit	Text Criticism	12. <i>Critical Apparatus</i>
textstructure	Default Text Structure	4. <i>Default Text Structure</i>
transcr	Transcription of Primary Sources	11. <i>Representation of Primary Sources</i>
verse	Verse	6. <i>Verse</i>

# Infrastruttura TEI

## Tabella Moduli TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

Component	Name	Example
Attribute Classes	att.*	att.global
Model Classes	model.*	model.biblPart
Macros	macro.*	macro paraContent
Datatypes	data.*	data.pointer

## Classi TEI

Le classi sono usate per esprimere due distinti tipi di **caratteristiche comuni** ad un insieme di elementi.

## Classi TEI

Gli elementi di una classe possono *condividere un insieme di attributi* oppure possono far *parte di uno stesso content model*.

## Classi TEI

- Un elemento appartenente ad una classe attributo condivide gli attributi con tutti gli altri elementi membri della stessa classe.
- Un elemento appartenente alla classe modello condivide il luogo del content model dove appare con gli altri elementi membri della stessa classe.

*In entrambi i casi un **elemento eredita** proprietà dalle classi di cui è membro.*



# Infrastruttura TEI

## Tabella Moduli TEI

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

class name	module name	see further
att.global.linking	linking	<i>16. Linking, Segmentation, and Alignment</i>
att.global.analytic	analysis	<i>17. Simple Analytic Mechanisms</i>
att.global.facs	transcr	<i>11.1. Digital Facsimiles</i>
att.global.change	transcr	<i>11.7. Identifying Changes and Revisions</i>

## Macro TEI

Le Macro sono shortcut per dichiarazioni che occorrono frequentemente.

Le Macro sono utilizzate in due modi diversi:

- per content model o parti di content model  
*frequently-encountered*
- per datatype di attributi

## Data Type TEI

I valori che possono assumere gli attributi sono definiti da tipi di dato all'interno delle *TEI datatype specification*.

## Data Type TEI

Le specifiche TEI definiscono i propri tipi di dato sfruttando altri tipi di dato primitivi e quelli derivati dalle specifiche W3C.

# Attributi Globali

Esempio @xml:lang

## xml:lang indica la lingua e il sistema di scrittura usato

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader xml:lang="en">
    <!-- ... -->
  </teiHeader>
  <text xml:lang="fr">
    <body>
      <div>
        <!-- chapter one is in French -->
      </div>
      <div xml:lang="de">
        <!-- chapter two is in German -->
      </div>
      <div>
        <!-- chapter three is French -->
      </div>
      <!-- ... -->
    </body>
  </text>
</TEI>
```

*Quasi tutti gli elementi TEI possono essere **classificati informalmente** come appartenenti alle seguenti categorie:*

### TEI element classification

- divisions
- chunks
- phrase-level elements
- inter-level elements
- components

### ■ **divisions**

Divisioni ad alto livello dei testi, molto spesso elementi annidati.

### ■ **chunks**

Elementi come i paragrafi e altri elementi simili i quali sono posizionati all'interno dei testi e divisioni. Solitamente non sono elementi che possono annidarsi o apparire all'interno di altri elementi di livello chunk.

### ■ **phrase-level elements**

Elementi che occorrono solo all'interno di elementi di livello chunk.

### ■ **inter-level elements**

Elementi che possono occorrere sia tra chunks all'interno di division, sia all'interno di essi.

### ■ **components**

Elementi che possono occorrere direttamente all'interno dei testi o delle divisioni di testo. E' una combinazione di elementi di livello inter e chunk.

# Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

## Struttura di un documento TEI

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



# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

5 Strumenti per edizioni critiche

6 Considerazioni Finali

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

**4 TEI: Codifica Apparato Critico**

5 Strumenti per edizioni critiche

6 Considerazioni Finali

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

**5 Strumenti per edizioni critiche**

6 Considerazioni Finali

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

5 Strumenti per edizioni critiche

6 Considerazioni Finali

# Progress status

Dalla Recensio  
all'Emendatio  
Digitale

A.M. Del  
Grosso

Presentazione

Panoramica  
Text Encoding  
Initiative

Panoramica  
eXtensible  
Markup  
Language

TEI: Codifica  
Apparato  
Critico

Strumenti per  
edizioni  
critiche

Considerazioni  
Finali

References

1 Presentazione

2 Panoramica Text Encoding Initiative

3 Panoramica eXtensible Markup Language

4 TEI: Codifica Apparato Critico

5 Strumenti per edizioni critiche

6 Considerazioni Finali