

Elementi Codifica TEI a.a. 2023-2024

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

CNR-ILC

`http://ilc.cnr.it/`

Istituto di Linguistica Computazionale “A. Zampolli”,
15th March 2024

Sommario della Lezione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

- 1 Introduzione al progetto TEI
- 2 Introduzione Text Encoding Initiative
- 3 Intestazione TEI
- 4 Elementi di base TEI
- 5 Infrastruttura TEI
- 6 Personalizzare TEI

Progress status

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI

Introduzione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Cos'è la TEI

la TEI - *acronimo di Text Encoding Initiative* - rappresenta un **punto di riferimento** per tutte le iniziative il cui scopo principale è quello di **digitalizzare risorse testuali** in ambito umanistico per fini di *ricerca* e di *conservazione*.

Introduzione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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**.

Per questo motivo il consorzio investe molte risorse per garantire la massima **accessibilità e divulgazione** delle tecnologie che da anni sviluppa.

Modularità della TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

- parleremo degli elementi basilari
 - Intestazione TEI (TEIHeader)
 - Elementi e attributi presenti in tutti i documenti TEI
 - Esempi di codifica
- parleremo del sistema Modulare della TEI

■ Moduli

■ Classi

■ Macro

■ Datatype

Modularità della TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

- parleremo degli elementi basilari
 - Intestazione TEI (TEIHeader)
 - Elementi e attributi presenti in tutti i documenti TEI
 - Esempi di codifica
- parleremo del sistema Modulare della TEI
 - Moduli
 - Classi
 - Macro
 - Datatype

I principi fondamentali della TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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 l’adottano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche svolte in seno alla comunità di riferimento (community-driven).

I principi fondamentali della TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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 l’adottano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche svolte in seno alla comunità di riferimento (community-driven).

I principi fondamentali della TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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 l’adottano per la creazione oppure per l’elaborazione dei documenti elettronici.
- La TEI cresce, matura, si evolve sulla base delle indicazioni e delle ricerche svolte in seno alla comunità di riferimento (community-driven).

Progress status

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 versione P4
- **2007:** nuova versione TEI P5, continuamente aggiornata

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

TEI Guidelines

Versione TEI P4

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

versioni P1 e P3 basate su SGML

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

TEI Guidelines: Obiettivi (dalle guide lines)

- 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 flexible framework into which unpredictable extensions can be fitted

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Che cosa offre la TEI

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

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Che cosa offre la TEI

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

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 dichiara 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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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: <http://www.tei-c.org/Lite/>)*

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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. Il tool ad oggi funzionante per la personalizzazione dello schema si chiama TEI ROMA
<https://roma.tei-c.org/>

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 (*One Document Does it all*).

TEI Roma

Il metodo da seguire per la versione P5

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

TEI Roma: Il formato ODD

tutorial: Getting Started with P5 ODDs

www.tei-c.org/Guidelines/Customization/odds.xml

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Il futuro della TEI

**negli ultimi anni l'approccio alla codifica TEI è stato
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 nativamente il markup di tipo stand-off (implementato un metodo ad-hoc)

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Il futuro della TEI

Importante: il formato XML non è l'unico formato per realizzare documenti TEI (in passato SGML).

TEI: Modello astratto

Se c'è una discrepanza fra Guidelines e schemi di codifica, la precedenza va alle Guidelines (**abstraction level**).

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Un documento TEI P5 'minimo'

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

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Caratteristiche del vocabolario TEI-XML

- gli **elementi TEI** rientrano nelle categorie generali di **elementi XML**
- 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/>`, `<milestone />`)
- gli elementi vuoti marcano una gerarchia differente

Intro Text Encoding Initiative

Gerarchie multiple

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<?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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<?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-<pagina n="6"/>
    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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<?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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<?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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

documento TEI - schema di intestazione TEI minima

Metadati essenziali riguardano 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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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>
```


Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Lo schema TEI permette di includere anche metadati registrati con altri schemi di codifica usando il descrittore `<xenodata>`

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
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>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 una 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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 `<emph>`
- **errori di interpretazione** del testo (che portano al tipo d'errore precedente o all'assenza del markup che andrebbe inserito)

Gli errori del primo tipo sono i più facili da individuare e correggere, quelli del terzo i più difficili

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>` (annidato in `<tagsDecl>`)

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI

Intro Text Encoding Initiative

Schemi di codifica TEI – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Contenuto del TEI header

Qualsiasi documento TEI deve includere una intestazione (TEIHeader).

TEI header: componenti

file description, encoding description, profile description,
revision description, no TEI metadata

Intro Text Encoding Initiative

Schemi di codifica TEI – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Testo in prosa

Molti elementi dell'intestazione contengono semplicemente del testo in prosa (giunti a qualche livello dell'albero del documento).

Testo in prosa

Alcuni elementi possono contenere testo in prosa suddivisa da **elementi di paragrafo** (*p*), oppure elementi più specifici strutturati, che a loro volta contengono testo in prosa.

Intro Text Encoding Initiative

Schemi di codifica TEI – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Grouping elements

Gli elementi che terminano con suffisso **Stmt** (esempio *titleStmt*), racchiudono un gruppo di elementi specifici e specializzati per la codifica di informazioni strutturate.

Grouping elements

In molti casi questi elementi possono contenere anche descrizioni con testo in prosa. Ciò permette al curatore della codifica di decidere se utilizzare una strategia strutturata, semi-strutturata o completamente destrutturata.

Intro Text Encoding Initiative

Schemi di codifica TEI – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Dichiarazioni

Gli elementi che terminano con il suffisso **Decl** (per esempio *refsDecl*) racchiudono informazioni inerenti specifiche prassi di codifica che sono state usate all'interno del testo digitale.

Dichiarazioni

Spesso queste prassi sono descritte in un linguaggio formale, attraverso una serie di dichiarazioni che identificano strutture complesse e descrizioni dettagliate.

*Le dichiarazioni possono essere riferite con l'attributo **decls** per ciascuna parte di testo a cui si applica.*

Intro Text Encoding Initiative

Schemi di codifica TEI – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Descrizioni

Gli elementi che terminano con il suffisso **Desc** (esempio *projectDesc*) contengono testo in prosa, che può essere anche strutturata usando specifici elementi.

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

documento TEI - schema di intestazione TEI minima

Metadati essenziali riguardano 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

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<?xml version="1.0" encoding="UTF-8" ?>
<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 – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Le altre componenti dell'intestazione TEI

■ elemento **xenoData**

Gli elementi di struttura specifici e specializzati che occorrono all'interno di questo elemento non sono parte del vocabolario TEI. Provengono da altri schemi di metadati come ad esempio il *Dublin Core*.

I metadati sono una componente essenziale di qualsiasi progetto di digitalizzazione

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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 – Intestazione

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

**Intestazione
TEI**

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Esempio Intestazione TEI raccomandata *Tratto dalle linee Guida della TEI. Capitolo 2 paragrafo 2.7*

Progress status

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI

Progress status

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI

Progress status

Elementi
Codifica TEI
a.a.
2023-2024

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

1 Introduzione al progetto TEI

2 Introduzione Text Encoding Initiative

3 Intestazione TEI

4 Elementi di base TEI

5 Infrastruttura TEI

6 Personalizzare TEI