

Elementi Codifica TEI a.a. 2021-2022

Angelo Mario Del Grosso

`angelo.delgrosso@ilc.cnr.it`

CNR-ILC

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

Istituto di Linguistica Computazionale “A. Zampolli”,
1st April 2022

Sommario della Lezione

Elementi
Codifica TEI
a.a.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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 (esiste una nuova versione beta)

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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

Intro Text Encoding Initiative

TEI

Elementi
Codifica TEI
a.a.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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 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

Elementi
Codifica TEI
a.a.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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 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

Elementi
Codifica TEI
a.a.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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.
2021-2022

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 – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Citazioni

- `<q>` testo citato da altre fonti: discorso diretto, esempi (nei dizionari), ecc.

Citazioni - Esempio

(La mia maestra della prima superiore mi salutò sulla porta della classe e mi disse: `<q rend='PRE mdash'>`Enrico, tu vai al piano di sopra, quest'anno; non ti vedrò nemmeno più passare!`</q>`)

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Citazioni

- `<quote>` frase o brano attribuito a fonte esterna

Citazioni - Esempio

```
(<p>E allora disse: <q rend=‘‘PRE lsquo POST  
rsquo’’>Ecco come comincia la Divina Commedia:  
<quote>Nel mezzo del cammin di nostra vita / Mi  
ritrovai per una selva oscura</quote>.</q></p>)
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Citazioni

- `<said>` testo pronunciato ad alta voce o pensato
- `<cit>` citazione con riferimento bibliografico

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Citazioni - Esempio

Lexicography has shown little sign of being affected by the work of followers of J.R. Firth, probably best summarized in his slogan, <cit>
<quote>You shall know a word by the company it keeps.</quote> <ref>(Firth, 1957)</ref> </cit>

Semplice riferimento bibliografico nell'esempio, possibile aggiungere un collegamento (a capitolo/sezione o a una specifica entrata bibliografica) usando l'attributo @target

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

TEI: elementi di citazione

A proposito di `<q>` e `<quote>` possono contenere non solo altri elementi simili (`<q>` e `<quote>`) ma anche elementi come `<p>`, `<l>`, ecc.:

Possibili problemi relativi alla gerarchia (tag overlap, gerarchie sovrapposte): `<p>Allora disse: <q>Sarò breve.</p><p>Ho finito.</q></p>`

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<p>
  <q>The Lord! The Lord! It is Sakya Muni
  himself,</q> the lama half sobbed; and under his
  breath began the wonderful Buddhist invocation:
  <q>
    <quote>
      <l>To Him the Way the Law Apart </l>
      <l>Whom Maya held beneath her heart</l>
      <l>Ananda's Lord the Bodhisat</l>
    </quote>
    And He is here! The Most Excellent Law is here
    also. My pilgrimage is well begun. And what
    work! What work!
  </q>
</p>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Nomi, numeri e date

- `<rs>` letteralmente “referring string”, nome o etichetta generica
`(<q>Mio caro <rs type=‘‘person’’>Filippo</rs></q>, gli disse <rs type=‘‘person’’>sua moglie</rs>.)`
- `<name>` nome proprio (di persona, luogo, ecc.)
`<q>Mio caro <name type=‘‘person’’>Filippo</name></q>`

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Nomi, numeri e date

- `<num>` un numero in qualsiasi formato `<num value='‘23’’>XXIII</num>`
- `<date>` una data in qualsiasi formato nato il `<date when='‘18680210’’>10 febb. 1868</date>`
- `<time>` l'ora in qualsiasi formato alle `<time when='‘8.00’’>otto del mattino</time>`

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Nomi, numeri e date

possibile usare un modulo specifico:

13 Names, Dates, People, and Places (<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/ND.html>)
(`<persName>`, `<forename>`, `<surname>`, `<roleName>`,
`<addName>`, ecc.)

Questo modulo permette una granularità molto maggiore grazie agli elementi specifici che mette a disposizione.

Possibile creare sistema prosopografico.

Ricchezza degli attributi (anche nella versione base)

La pagina del titolo

elemento <titlePage> può contenere:

- <docTitle> e (o direttamente) <titlePart>: titolo anche in più parti
- <docEdition>: informazioni riguardo l'edizione
- <byline> e (o direttamente) <docAuthor>: autore
- <docImprint>: informazioni di stampa, a sua volta include

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

La pagina del titolo

elemento `<docImprint>` può contenere:

- `<publisher>`: editore
- `<pubPlace>`: luogo di stampa
- `<date>`: data
- `<docDate>`: data

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<front>
...
  <titlePage>
    <titlePart>Cuore</titlePart>
    <byline>di <docAuthor>E. de Amicis</docAuthor>
    </byline>
    <docEdition>Edizione integrale</docEdition>
    <docImprint>
      <publisher>Newton Compton editori</publisher>
      <pubPlace>Roma</pubPlace>
      <date>1994</date>
    </docImprint>
  </titlePage>
</front>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<titlePage>
  ...
  <docAuthor>E. DE AMICIS</docAuthor>
  <docTitle>
    <titlePart type="main">CUORE</titlePart>
    <titlePart>Libro per i ragazzi</titlePart>
  </docTitle>
  <docEdition>98.a edizione</docEdition>
  <graphic url="publisher.png" />
  <docImprint>
    <pubPlace>MILANO</pubPlace>
    <publisher>FRATELLI TREVES, EDITORI</publisher>
    <date>1889</date>
  </docImprint>
</titlePage>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Collegamenti interni ed esterni

- `<ptr/>` specifica un puntatore a un altro *luogo* (un altro punto dello stesso testo o di un altro testo)

`<p>`Il sistema di puntatori è basato sul meccanismo W3C Xpointer. Per maggiori informazioni si veda il capitolo `<ptr target="#SAid"/>`; per le specifiche si veda `<ptr target='http://www.w3.org/TR/xptrxpointer/' />`; anche `<ptr type="image" target="#fig22"/>`.`</p>`

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Collegamenti interni ed esterni

- `<ref>` specifica un puntatore a un altro *luogo*, può includere del testo.

Si veda il `<ref>`terzo capitolo, p. 24`</ref>`. Si veda il `<ref target='‘#cap3.24’’>`terzo capitolo, par. 24`</ref>`

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Liste e tabelle

- `<list>` qualsiasi tipo di lista (type per specificare)
- `<head>` intestazione (titolo) della lista
- `<item>` un elemento della lista
- `<label>` numero o esponente associato all'`<item>`
- `<headLabel>` intestazione per gli esponenti della lista
- `<headItem>` intestazione per gli elementi della lista

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Agli item può essere associata o meno una etichetta (label), ma in caso affermativo deve essere presente per tutti.

Liste e tabelle

```
<list><head>Ingredienti:</head> <item>un cucchiaino  
di zucchero;</item> <item>mezzo chilo di  
farina;</item> <item>due uova.</item></list>
```

*Usare l'attributo type per specificare il tipo di lista.
Il testo non deve essere necessariamente ordinato come lista
per essere marcato come tale*

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<list type="simple">
  ...
  <head>Lista della spesa:</head>
  <item>pane;</item>
  <item>frutta;</item>
  <item>verdura;</item>
  <item>latte;</item>
  <item>farina;</item>
  <item>uova;</item>
  <item>tovaglioli;</item>
  <item>bicchieri;</item>
  <item>piatti.</item>
</list>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<list type="gloss">
...
  <head>Tecnologie XML:</head>
  <label>XSL</label>
  <item>eXtensible Stylesheet Language</item>
  <label>XSLT</label>
  <item>XSL Transformations</item>
  <label>XSLFO</label>
  <item>XSL Formatting Objects</item>
  <label>XQuery</label>
  <item>XML Query Language</item>
  <label>XPath</label>
  <item>XML Path Language</item>
</list>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Tabelle

gli elementi per tabelle sono `<table>`, `<row>`,
`<cell>` disponibili con il modulo figures, v. il
cap. 14 Tables, Formulae and Graphics
(<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/FT.html>)

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<table rows="3" cols="2">
...
  <row>
    <cell>HTML</cell>
    <cell>
      Derivato da SGML (come applicazione),
      il linguaggio del World Wide
      Web.</cell>
    </row>
    <row>
      <cell>XML</cell>
      <cell>Derivato da SGML (per semplificazione).</cell>
    </row>
    <row>
      <cell>XSLT</cell>
      <cell>Fogli di stile per XML.</cell>
    </row>
  </table>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Note

- l'elemento `<note>` permette di inserire una nota di qualsiasi tipo (attributi *type* e *place*)
- la nota può appartenere al testo codificato, o può essere opera di chi lo codifica (attributo *resp*)
- se non nel flusso del testo usare gli attributi *target* e *targetEnd* per stabilire collegamento preciso

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Note - Esempio

Le <title>Guidelines</title><note n="2"
place="foot" resp='‘amdg’>C.M. SperbergMcQueen
and Lou Burnard, <title>Guidelines for Electronic
Text Encoding and Interchange</title> (Chicago,
Oxford: Text Encoding initiative, 2002).</note>
sono nate con lo scopo di ...

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Esercizio

Utilizzare puntatori, note, lista, glossario marcare nomi, rs, term.

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Bibliografia

- `<bibl>` citazione bibliografica di tipo *flessibile*: possono essere presenti solo alcuni elementi; usata anche nel corpo del testo
- `<biblStruct>` citazione bibliografica di tipo strutturato: gli elementi devono seguire uno schema fisso
- `<biblFull>` citazione bibliografica di tipo strutturato completo: gli elementi devono seguire uno schema fisso e devono essere tutti presenti
- `<listBibl>` lista di citazioni bibliografiche in uno qualsiasi dei formati citati sopra

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Una voce bibliografica può essere codificata impiegando tre possibili tipi:

Tipi di pubblicazione

- `<analytic>` “analitico”: informazioni su un titolo che non costituisce una pubblicazione autonoma (articolo di rivista o in una miscellanea)
- `<monogr>` “monografico”: informazioni su una pubblicazione, anche di tipo periodico (rivista)
- `<series>` “serie”: informazioni su di una serie editoriale

tipicamente impiegati con gli elementi `<biblStruct>` e `<biblFull>`, più flessibilità con `<bibl>`

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Titolo, autore, curatore

- `<title>` il titolo della pubblicazione; più livelli (attr. `level`)
- `level="a"` articolo (article)
- `level="j"` rivista (journal)
- `level="m"` monografia (monograph)
- `level="s"` serie (series)
- `level="u"` non pubblicato (unpublished)

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Titolo, autore, curatore

- `<author>` autore della pubblicazione
- `<editor>` curatore della pubblicazione

*Questi elementi possono contenere direttamente il testo con il nome oppure elementi più complessi (`<persName>` e *sub-elementi*)*

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Elementi descrittivi

- `<imprint>` raggruppa informazioni relative alla stampa o diffusione di una pubblicazione
- `<publisher>` l'editore responsabile della produzione a stampa o diffusione
- `<pubPlace>` il luogo di pubblicazione
- `<date>` la data di pubblicazione

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Elementi descrittivi

- `<biblScope>` l'estensione (in pagine, volumi) della pubblicazione;
- con attributo *unit* permette di specificare il tipo di estensione ad es. se si tratta di pagine
- `<citedRange>` porzione di lavoro citato.

`<biblScope>` può essere usato anche all'interno di
`<series>`

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<biblStruct xml:id="Robinson1993d">
  ...
  <monogr>
    <author>Robinson, P.</author>
    <title level="m">The Digitization of Primary
    Textual Sources</title>
    <imprint>
      <pubPlace>Oxford</pubPlace>
      <publisher>Office for Humanities
      Communication</publisher>
      <date>1993</date>
    </imprint>
  </monogr>
</biblStruct>
```


Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<biblStruct xml:id="Mohler2001">
  ...
  <analytic>
    <author>Mohler, Peter Ph.</author>
    <author>Zuell, Cornelia</author>
    <title level="a">Applied Text Theory:
      Qualitative Analysis of Answers to Open Ended Questions
    </title>
  </analytic>
  <monogr>
    <editor>Mark D. West</editor>
    <title level="m">Application of Computer Content Analysis</title>
    <imprint>
      <pubPlace>Westport CN</pubPlace>
      <publisher>Ablex</publisher>
      <date>2001</date>
      <biblScope unit="pages">116</biblScope>
    </imprint>
  </monogr>
</biblStruct>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<biblStruct xml:id="Ester1994">
  ...
  <analytic>
    <author>Ester, Michael</author>
    <title level="a">Digital Images in the Context of
Visual Collections and Scholarship</title>
  </analytic>
  <monogr>
    <title level="j">Visual Resources</title>
    <imprint><date>1994</date></imprint>
    <biblScope unit="vol">X,1</biblScope>
    <biblScope unit="pages">23</biblScope>
  </monogr>
  <note type="key">iproc</note>
</biblStruct>
```

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Componenti non testuali: immagini

- in formato XML (*SVG, il formato TEI*)
- in formato binario (*TIFF, JPEG, PNG*)

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Componenti non testuali: immagini

- `<graphic/>` contiene un puntatore a una immagine
- attributo **url** può essere locale o esterno
- `<binaryObject>` dati binari che costituiscono una immagine
- `<media>` riferimento a file audio, video ecc.

Intro Text Encoding Initiative

Schemi di codifica TEI: Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Componenti non testuali: immagini

- `<figure>` raggruppa informazioni relative a un'immagine
- `<head>` titolo o didascalia relativa all'immagine
- `<figDesc>` descrizione dell'immagine
- `<graphic>` puntatore a un'immagine

`<figure>` nella P5 fa parte del modulo figures
(<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/FT.html#FTGRA>)

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Nella TEI P4 bisogna inserire dichiarazioni per le immagini nell'intestazione

```
<!DOCTYPE figEx [  
  <!NOTATION JPEG PUBLIC 'ISO DIS 10918//NOTATION  
    JPEG Graphics Format//EN'>  
  <!ENTITY PDatabase  
    SYSTEM 'support/PDatabase.jpeg' NDATA JPEG> ]>  
<figEx>  
  <p>The system is build around a relational  
    database, called the <soCalled>palaeographical  
    database</soCalled> ... </p>  
  <figure entity="PDatabase">  
    <figDesc>Schematic of the SPI system.</figDesc>  
  </figure>  
</figEx>
```

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Immagine in un documento P5

```
<p>Questo fenomeno è chiaramente visibile nella  
figura che segue: <graphic url="diagramma.png"/>  
Altre variazioni delle crescite sono  
disponibili...</p>  
<head> <graphic  
url="http://www.name.org/hpral02.gif"/> </head>
```

immagine impiegata come intestazione in un documento P5

Intro Text Encoding Initiative

Schemi di codifica TEI – Moduli base

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Esercizi documenti di tipo e genere diversi

- codificare usando gli opportuni elementi TEI un articolo di rivista
- codificare usando gli opportuni elementi TEI una cartolina

Progress status

Elementi
Codifica TEI
a.a.
2021-2022

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

TEI framework

La tecnologia TEI ha un framework concettuale diviso in

- Moduli
- Classi
- Macro
- Tipi di Dato

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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, Formulae, 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

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

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.

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.

Progress status

Elementi
Codifica TEI
a.a.
2021-2022

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 – Personalizzare TEI

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

Personalizzare TEI

Nessun progetto di codifica richiede di utilizzare tutte le specifiche definite dalle linee guida della TEI.

Personalizzare TEI

La TEI fornisce un insieme specifico di elementi che può essere usato per creare uno schema TEI puntuale e ritagliato sulle specifiche del progetto di codifica in corso.

Intro Text Encoding Initiative

Schemi di codifica TEI – Personalizzare TEI

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI

```
<schemaSpec ident="tei-custom"
...
  start="TEI teiCorpus" >
    <moduleRef key="analysis"
      include="interp interpGrp pc s w" />
    <moduleRef key="linking"
      include="anchor seg" />
    <moduleRef key="tei" />
  </schemaSpec>
```

*Si selezionano o si escludono gli elementi attraverso gli attributi
@include e **@exclude***

Intro Text Encoding Initiative

Schemi di codifica TEI – Personalizzare TEI

Elementi
Codifica TEI
a.a.
2021-2022

A.M. Del
Grosso

Modifica agli elementi e attributi dei Moduli

```
<classSpec type="atts" ident="att.dataable.w3c"
...
module="tei" mode="change" >
  <attList>
    <attDef ident="notAfter" mode="delete" />
    <attDef ident="from" mode="delete" />
    <attDef ident="to" mode="delete" />
  </attList>
</classSpec>
```

Introduzione
al progetto
TEI

Introduzione
Text Encoding
Initiative

Intestazione
TEI

Elementi di
base TEI

Infrastruttura
TEI

Personalizzare
TEI