

Per favore, non copiatela.



La per nulla professionale (ma immediata) guida TEL di Stefano Rizzo.



New! Sexier! Improved (kinda)!

Nella speranza che vi sia utile.

Ho davanti un testo da codificare.

Apro Oxygen XML Editor 22.1

(Leggenda narra che questa guida funzioni anche con altre versioni, ma io declino ogni obbligo legale. Sono povero.)

Regola generale fondamentale: ogni tag che si apre, si chiude: <ciò che si apre> </ciò che si apre>. Per comodità nella guida non chiudo niente se non negli esempi, che mostrano sempre una parte interessata; inoltre, i termini sono messi tutti in minuscolo, mentre per esempio si dovrebbe scrivere <teiHeader>, ma tanto Oxygen segnala cose così e tende al completamente automatico mentre scriviamo. Sbagli così aiutano a prendere dimestichezza con i termini corretti. Sfortunatamente, non tutti i programmi mettono a disposizione questa funzione.

Si incolonna per gerarchia. Si inarca fino alla fine della sezione.

```
<teiheader>
```

```
</teiheader>
```

```
<text>
```

```
    <body>
```

```
    </body>
```

```
</text>
```

La forma corretta di una codifica è

<codifica attributo="descrizione nostra non sempre completata automaticamente"

attributo2="più attributi possono essere inseriti" attributo3="mi raccomando le virgolette"> il testo che è oggetto di codifica **</codifica>**

I colori nel programma sono diversi in realtà, ma poco importa se si capisce come funziona. Avrete notato che non ci sono spazi dentro gli attributi che noi scriviamo: solo in occasioni o in codifiche speciali è consentito mettere lo spazio, per tutto il resto si usa l'underscore (ossia il trattino basso: _), e se siete qui a leggere questa guida difficilmente avrete a che fare con occasioni simili.

Certi elementi della codifica vanno inseriti, come minimo. Sono pertanto segnati come **MINIMO**.

Due funzioni molto carine: invece di scrivere sempre ogni parte della codifica, la combinazione di tasti CTRL + E consente di inserire una codifica nella linea di testo considerata; se abbiamo dubbi circa la visualizzazione del testo una volta finita la codifica (ad esempio in parti di testo marcate in corsivo o grassetto), possiamo visualizzare il nostro prodotto finito selezionando in basso a sinistra la voce "Author" (l'ultima tra *text/grid/author*). *Sonia sei il numero uno per aver trovato questo life hack.*

Se negli esempi o nelle codifiche troviamo la formuletta "**&mdash**" oppure "**—**" non abbiate panico, è solo un modo per codificare caratteri speciali. Nello specifico, veniva usato per il trattino lungo (—) tipico dei dialoghi in forma diretta, di un'espressione di enfasi oppure di una cesura. Stessa cosa accade con **&ndash** e simili.

Salvo rare occasioni, come nella descrizione della lingua <language> (vedi <profiledesc>), **bisogna sempre mettere in un <p> una nostra descrizione per esteso: <availability status="free"> <p> Questo testo è disponibile online </p> </availability>**.

In genere la codifica è preimpostata nelle parti più universali.

Non mi chiedo ancora nulla

ma

Mi chiedo se sia un corpus:

<teicorpus.[numero o nome del corpus qui]> se ci sono dei corpora vanno specificati ogni volta per numero, un **ATTRIBUTO version** potrebbe aiutare con il lavoro futuro

Non mi chiedo ancora nulla del testo: teiHeader

<teiheader> MINIMO intestazione della TEI

<filedesc> MINIMO

<titlestmt> MINIMO, titolo opera + proprietà intellettuale (che cosa? chi?)

<title> titolo

<author> autore. È bene attribuire qui un **xml:id** = "primetrelletteredelnome" al quale poter fare riferimento con **ATTRIBUTO ref** durante eventuali episodi di analisi. Ovviamente si possono mettere anche i nomi per esteso, ma le prime tre lettere del nome di solito bastano. Ed è comodo.

<sponsor> organizzazione o istituzione sponsor

<funder> finanziatore della codifica

<principal> capoprogetto della digitalizzazione (tu, se lo stai digitalizzando per primo, Di Silvestro se sei un buontempone), in generale si potrebbe saltare in favore di <respstmt>

<respstmt> responsabilità di ulteriori quando gli elementi dei curatori e autori non bastano

<resp> responsabilità della persona *ad esempio, "Edizione elettronica a cura di"*

<name> nome della persona di cui sopra, ossia quantomeno il tuo nel primo <resp>.

<editionstmt> edizione del testo che stiamo digitalizzando (quale quando?)

<edition> particolare edizione

<raspstmt> dichiarazioni ulteriori se tutto il resto non basta

<extent> info sulle dimensioni della codifica (e) sul supporto di memoria (quanto pesa? es "117Kb in UTF-8")

<publicationstmt> **MINIMO**, info su distribuzione del testo

<publisher> organizzazione responsabile della pubblicazione del testo

<distributor> responsabile o organizzazione responsabile della distribuzione del documento

<authority> responsabile della concessione alla messa a digitale del documento se non coincide con i campi sopra

questi tre possono avere sotto:

<pubplace> luogo di pubblicazione di un documento

<address> indirizzo

<idno> identificatore di unità bibliografica, può essere **TYPE** (codice ISBN o simili)

<availability> info sulla disponibilità di un testo ed esempio copyright, può essere **STATUS** (*limitato/libero/parziale*)

<date> una data

<seriesstmt> informazioni sulla collana (fa parte di una collana? quale?)

<title> nome collana

<idno> identificatore (*ad esempio ISBN*)

<respstmt> autore della collana

<notesstmt> informazioni non altrimenti reperibili nella codifica (c'è altro da sapere?)

contiene uno o più sottoelementi

<note> nota

<sourcedesc> **MINIMO**, fonte del testo trascritto (da dove è stato preso?)

<bibl> riferimento bibliografico non strutturato. Se la fonte originale non corrisponde a quella dalla quale abbiamo preso il testo, possiamo codificare diversi **type** di fonte (= "digitale" e poi "originale", *ad esempio*). Possiamo fare una cosa più carina articolando con sotto:

<title> al quale si può inserire un **ATTRIBUTO level**, per indicare se il testo è una monografia con = "m", se è un elemento analitico (*articolo, poesia, lavoro pubblicato come parte di elemento più grande*) con = "un", se è una rivista con = "j", se è una serie o una raccolta con = "s", se è un inedito con = "u".

<author> autore

<editor> editore (*ed esempio type="revisore" o "traduttore"*)

<date> data

<publisher> casa editrice

<bibfull> riferimento bibliografico rigorosamente strutturato. Molto complesso, è più fluido articolare noi bene il <bibl>.

<listbibl> lista di citazioni bibliografiche

Ora troviamo altre cose fuori <filedesc> ma dentro la <teiheader>

<encodingdesc> metodi e editoriali della trascrizione e della codifica

può essere in prosa oppure:

<projectdesc> il fine della codifica più info di produzione

<samplingdecl> principi della selezione dei testi del corpus o della selezione

<editoriadescl> principi e pratiche editoriali:

<p> su correzioni apportate al testo

<p> su normalizzazione alla fonte

<p> su trattamento virgolette (delle citazioni) rispetto a originale

<p> su segmentazione testo (*es frasi, unità toniche, strati grafemici...*)

<p> su interpretazione o interpretazione analitica

<tagsdecl> info su marcatori ad un testo XML

Può essere composta usando

<tagusage> info su uno specifico elemento dentro <text> con **ATTRIBUTI**: gi per identificatore generico, **occurs** per il numero di occorrenze; oppure con l'attributo **render** per fornire all'identificatore <rendition> (qua sotto) come tale elemento vada reso graficamente

<rendition> info su presentazione grafica di uno o più elementi

<refsdecl> modo di costruzione dei sistemi di riferimento canonici del testo, in prosa.

<classdecl> tassonomia dell'insieme dei codici di classificazione.

Richiede almeno uno tra:

<taxonomy> tipologia per classificare testi implicitamente (*citazione bibliografica*) o esplicitamente (*tassonomia strutturata*)

<bibl> citazione bibliografica non strutturata, le componenti possono essere anche esplicitamente etichettate

<category> singola categoria descrittiva

<catdesc> descrizione in prosa della singola categoria

<profiledesc> metainformazioni del testo:

<creation> info su creazione

<langusage> lingue, sottolingue, registri, dialetti in un testo. È possibile allegare un **ATTRIBUTO** identificativo **ident**. Non va messa la <p>. *Un esempio per un testo in italiano è **<langusage ident="ita">Italiano</langusage>***

<textclass> natura/temi di un testo come classificazioni, thesaurus o simili

<keywords> parole e temi chiave, con possibilità di **ATTRIBUTO** **scheme** che identifica il vocabolario dell'insieme di parole e che si può eventualmente collegare al sistema di classificazione in <taxonomy>

<catref> collegamento tra testo e categoria di una tassonomia, l'**ATTRIBUTO** **target** identifica le connessioni

<revisiondesc> diario con liste <change> dei cambiamenti sull'elaborato elettronico

<change> elemento con dentro:

<date> data del cambiamento

<respstmt> responsabilità, qualora mancasse altrove

<name> chi modifica

<item> componente di una lista

Mi chiedo che elementi abbia il testo:

<text> **MINIMO**

<front> peritesto iniziale

<group> raggruppa testi o gruppi di testi

<body> **MINIMO** intero corpo del testo

<back> peritesto finale

Ora la carta jolly importante: i **<div>**. Questi indicano una divisione qualsiasi del testo, può essere più che altro utile a noi per distinguere un livello del testo al quale dare una numerazione o un titolo. Tralasciando l'eventuale utilità nostra, è utile a prescindere da qualsiasi cosa si analizzi. Ad esempio, dopo il <text>, prima del <body>: **<div type="capitolo" n="1"><head rend="bold"> <p> Qui metto il titolo </p> </head>**. Dopo il body, la chiusura della sezione </div>. I div possono essere numerati anche internamente, ad esempio **<div1>**.

Mi chiedo che tipo di <text> io abbia davanti:

È una prosa?

È una poesia?

È una scenografia?

- È una **prosa**: i fondamentali. Mi concentro su tutto ciò che si possa codificare come base e mi fa da scheletro per i successivi casi. Indico le divisioni logiche del testo con **<div>** numerandolo per volta (**<div1>...**), metto **<p>** al paragrafo, **<q>** per il discorso diretto o una citazione. Un piccolo appunto: dentro **<q>** vanno messi anche gli indicatori del discorso diretto; se non facciamo così o se li sostituiamo dobbiamo rendere conto di questo nella **<encodingdesc>**.

Se la pagina si interrompe, usare nella parte interessata l'elemento particolare **<pb/>** (si nota che non inizia e non ha fine per la stanga a destra, non alla solita posizione). Se invece vogliamo prestare attenzione a codificare ogni singola pagina, usiamo **<pb>** con dentro l'**ATTRIBUTO** **n="numerodipagina"** e, parte fondamentale, una sbarra alla fine: **<pb n="1"/>**, così che non ci sia bisogno di chiudere successivamente la pagina. Altrimenti, possiamo anche usare **<pb n="pagina">** e chiuderlo successivamente con un **</pb>**, ma perché farlo se abbiamo già un setup così efficace?

Se ci interessa evidenziare l'inizio di una riga tipografica si segna con **<lb>**. Ambo **<pb>** e **<lb>** possono avere l'**ATTRIBUTO** **ed** per specificare l'edizione. In base a cosa poi io cerco nel testo, possiamo avere anche **<persname>** per indicare il nome del personaggio o **<placename>** per i nomi dei luoghi. Se una parte del testo ha una formattazione particolare, con **<hi>** o -meglio- con **<seg>** (vedremo meglio sotto) e l'**ATTRIBUTO** **rend** si può segnare questa peculiarità (ad esempio, **rend="bold"** è grassetto, **rend="italic"** è corsivo).

Ogni volta che si codifica un discorso diretto o indiretto, bisogna usare le stesse regole che si impiegano nella sezione, più avanti, sulla codifica teatrale ma dentro **<q>**: **who**, **towhom**... Ed infine al posto **<castlist>** ma con le medesime regole di articolazione, usiamo **<listperson>** per tutti i personaggi.

<body>

<p><q>— Perché hai deviato? —</q> domandò Ivan. <q>— Perché ora ti sei fermato? Cosa guardi? Quella casa? Perché ti interessi a quella casa? </q></p>

</body>

- È una **poesia**: mi dovrebbero interessare quantomeno la metrica e le figure retoriche. Codifico i gruppi di versi (*strofa*, *refrain*, *paragrafo in versi*) con **<lg>**, poi codifico le righe singole di poesia con **<l>** che può avere **ATTRIBUTI part** per specificare se il verso è completo o meno accompagnato da **f** (parte finale di verso incompleto) / **y** (verso incompleto metricamente) / **n** (non lo so se è completo o no) / **i** (parte iniziale di verso incompleto) / **m** (parte mediana di verso incompleto). Posso segnare la metrica dentro **<lg>** o in **<l>** in base alle eventuali variazioni, con l'**ATTRIBUTO** **met="dodecasillabo"** o simili. È importante più che altro numerare i versi con l'ATTRIBUTO n="numerodelverso". **<rhyme>** indica una rima, possiamo inserire in **<lg>** lo schema rimico tra virgolette ed in maiuscolo, e nei versi possiamo specificare la lettera (sempre in maiuscolo) dello schema rimico corrispondente alla parola rimante con l'**ATTRIBUTO** **label**, come viene mostrato nel secondo esempio.

```

<body>

  <div type="composizione"> <head rend="bold"> <p> Mattino </p> </head>

    <lg type="doppio verso" met="quadrisillabo" rhyme="none">

      <l n="1"> M'illumino d'immenso </l>

    </gl>

    <gl>

      <l n="2"> Questo verso è la famosa seconda stanza di "Mattina" di Ungaretti </l>

    </gl>

  </div>

</body>

```

Esempio con le rime:

```

<lg rhyme="AA">

  <l n="1"> Di poesia, quest'<rhyme label= A> esempio </rhyme></l>
  <l n="2"> è proprio uno <rhyme label= A> scempio </rhyme> </l>
  <l n="3"> ma se siete <rhyme label =B> disperati </rhyme> </l>
  <l n="4"> almeno qui sarete <rhyme label = B> accompagnati </rhyme> </l>

</lg>

```

Oh-oh, che sorpresa: la poesia ha delle **figure retoriche**. Assurdo, vero? Ma come facciamo a codificarle?

Di base un metodo efficace per imparare a codificarle è il seguente: apro e chiudo un **<interp>** nella parte di testo con la figura retorica, inserisco dentro l'**ATTRIBUTO ana** con un richiamo **"#nomedellafiiguraretorica"**. Successivamente, nel **<body>**, creiamo un **<interpgrp="figure retoriche">** con sotto, per ogni tipo di figura retorica, un **<interp xml:id="nomedellafiiguraretorica">** *nome per esteso della figura retorica e versi nella quale compare in formato v(numero)* **</interp>**. Nota bene che ciò che va dopo un hashtag (e tra le virgolette) deve essere uguale a quello che è tra virgolette nell'**xml:id**.

Ecco un esempio:

```

<body>

  <l n="1">La nebbia a gl'<interp ana="#metafora">irti</interp> colli</l>

</body>

<back>

  <interpGrp>

    <interp xml:id="metafora">metafora v(1)</interp>

  </interpGrp>

</back>

```

Simone sei un mito per averlo capito

Questo metodo ha però un grosso problema: poniamo il caso di una figura retorica

articolata o di più figure retoriche dentro lo stesso segmento di testo (si pensi ad esempio, un ossimoro dentro un'assonanza, o una metafora dentro una perifrasi), non possiamo assegnare più xml:id, né "incassare" una codifica dentro l'altra con più xml:id. E questo è male, perché se vogliamo sapere quante figure retoriche ci sono nel componimento, finiremmo per calcolare in modo diverso o per lasciare indietro qualcosa a causa della struttura. Mi spiego peggio: c'è un testo con 15 assonanze, e codifichiamo tutte le assonanze con xml:id="assonanza" assegnato all'<interp>. Un'assonanza coincide però con una climax, ma non possiamo dividerle usando l'interp, pertanto siamo costretti ad assegnare un id diverso a questo caso, ad esempio xml:id="assonanzapiùclimax". Se chiediamo a un programma di contare, le assonanze scendono da 15 a 14, perché la macchina non può sapere di quel caso particolare: l'id non coincide.

Come ovviare?

Ci sono un paio di modi, come giocare con l'attributo <seg> per evidenziare un segmento con il suo rispettivo ana="figurainquestione", ma la cosa migliore è chiedere al proprio Responsabile. Sia questi un professore o un project manager.

Più scorrevoli sono gli enjambement, che rientrano più che altro nella composizione del verso invece che nella sezione delle figure retoriche: dentro <l> possiamo inserire l'**ATTRIBUTO enjambement** per indicare che alla fine del verso è presente un enjambement, che solitamente sarà ="str" in quanto di tipo forte, ma ci possono essere altri attributi (y per sì, n per no, *strong* per forte e *weak* per lieve). Aggiungiamo dentro <l> l'**ATTRIBUTO ana=#enjambement** (o acciordiatelo, se volete, basta che sia identificato bene nell'<interpgrp> del <back>).

- È una **scenografia**: si presta attenzione alla battuta, a chi la dice, al che contesto nella quale viene detta. Si sa <sp> per la battuta con **ATTRIBUTO who** che identifica via ID il parlante ed eventualmente **towhom** per la persona alla quale il messaggio è rivolto, <speaker> che fornisce i nomi dei parlanti del testo, <stage> che contiene le direttive di scena con **ATTRIBUTI type** che noi definiamo (*entrata, uscita, scenografia, dizione...*).

<body>

<sp who="GUB"> <speaker> Serafino Gubbio operatore </speaker> <p>— Che vuole che le dica? Lei in questo momento, lo riconoscerà, è molto eccitato. </p> </sp>

<sp who="FER"> <speaker> Carlo Ferro </speaker> <p>— Ma posso esser calmo? </p> </sp>

</body>

La lista dei personaggi coinvolti va segnata o nella sezione <front> o <body> con l'elemento <castlist>. Questo ha parecchi sottoelementi, quindi viene fornito un esempio con commento a lato:

<castList>

<castGroup>

Gruppo di personaggi

<head>Mendicanti</head>

Nome del gruppo

<castItem>

Singolo elemento del gruppo

<role>Aafaa</role>

Ruolo

<actor>Femi Johnson</actor>

Attore

</castItem>
<castItem>
</castGroup>
</castList>

Il grosso è fatto, ma ho dei peritesti.
Torno al lavoro.

Peritesti iniziali.

Il frontespizio:

<titlepage>

<doctitle> per il titolo del documento, diviso in

<titlepart> per le suddivisioni delle parti dell'opera, con **ATTRIBUTI type** (es principale, sottotitolo, descrittivo, alternativo...)

<byline> dichiarazione di responsabilità (si può trovare anche alla fine, non nel frontespizio)

<docauthor> autore della digitalizzazione. Spesso si trova sotto <byline> e non come incolonnato qui

<docdate> data di edizione del documento

<docedition> edizione del documento

<docimprint> dichiarazione delle note tipografico-editoriali

<epigraph> citazione all'inizio di una sezione (qui frontespizio, ma si può usare per citazioni all'inizio di un capitolo)

I materiali introduttivi.

Diamo agli elementi <div> (vedi sopra, in "prosa") i **type** *prefazione, introduzione, dedica, abstract, ringraziamenti, sommario e decorazione*.

SOLO PER LE EPISTOLE troviamo in più gli elementi <salute> (saluto), <signed> (formula di commiato), <byline> responsabilità dell'epistola, <dateline> per descrivere luogo/data o simili, <argument> per una descrizione in prosa dei temi, <opener> per raggruppare date/firme/saluti/espressioni all'inizio e <closer> per queste ultime ma poste alla fine.

Peritesti finali.

sempre con <div>, troviamo i **type** *appendice, glossario, note* con ogni nota codificata da <note> (vedi sotto), *bibliografia* codificata con una lista (vedi sotto) speciale <listbibl> i cui elementi singoli sono <bibl> (vedi sotto), *indice* codificato come lista strutturata, infine *colophon*.

Mi chiedo cosa mi possa interessare:

perché mi frega quel che mi frega?

Gli attributi sono tutto nella codifica TEI. Avere molti marcatori vuoti serve a poco: bisogna specificare di cosa si stia marcando un aspetto per renderlo ricercabile o utile alla codifica.

Ad esempio troviamo per molti marcatori attributi **type** per spiegare qualcosa (es. **<div type="chapter">** per dire che questo blocco è un capitolo);

n per numerare (es. **<p n="1">** per la prima frase);

xml:id per identificare univocamente un tag (es. **<div xml:id="c1p1">**);

who e **thowhom** per dire chi dice qualcosa a chi (es. **<q who="#GUB" towhom="#FER">**);

ref rimanda a un riferimento/risorsa (**<persname ref="Lia">**)

In realtà ce ne sono molti di più. Ma tutto dipende sempre da cosa cerco, cosa voglio che sia ricercabile. Ciotti cita **rend** per segnalare una resa grafica del testo (*corsivo, grassetto, sbarrato...*).

Prima degli attributi ci sono però i marcatori. Stesso discorso di prima. Vuoi parlare del titolo? Marcatore **<title>** e **ATTRIBUTI level** (=m se sono monografici, =s per una collana, =j per un periodico, =u per inedito, =a per analitico) e **type** se è un titolo di tipo particolare (*principale, subordinato, parallelo* a.k.a. "in altra lingua", *abbreviato*).

Problema: **non so perché** mi interessa questa parte di testo. Soluzione: piazzaci un **<hi>** per marcarlo ma sospendendo il giudizio su di essa, lasciando che altri se ne occupino.

Se invece **lo sai**, puoi descriverne l'importanza con **<emph>** per sottolineare un'enfasi, **<foreign>** se è in una lingua diversa con **lang** a specificare quale, **<mentioned>** se è una citazione, **<term>** se è un tecnicismo. E così via. È "tutto" qui, basta usare la barra di ricerca: [The TEI Guidelines \(tei-c.org\)](http://The-TEI-Guidelines.tei-c.org).

Un indicatore simile a **<hi>** e dalla stessa flessibilità di uso è **<seg>**, che si usa per parlare/descrivere/spiegare un segmento di testo. Si possono attribuire una marea di attributi utili e sbarazzini a **<seg>**.

Le **citazioni** in effetti sono particolari: essendo un testo codificato per avere rimandi, oltre a **<mentioned>** possiamo trovare **<socalled>** per una parola della quale l'autore non si fa carico (tipo le parole tra virgolette o in corsivo in una narrazione normale), **<gloss>** per una glossa che può avere -quindi- un **target** che identifica ciò a cui la glossa fa riferimento.

Le **citazioni bibliografiche** sono diverse, si codificano con **<bibl>** e molteplici altri elementi **<author>**, **<biblscope>** per l'estensione (numeri di pagina, nome di un sottotesto di un'opera più ampia), **<date>**, **<editor>** con **role**, **<imprint>** per raggruppare un'unità bibliografica, **<publisher>**, **<pubplace>**, **<series>** se è una collana ed, infine, ovviamente, **<title>** con **type** e **level** per il livello bibliografico.

Altre cose che possono richiamare la nostra cura sono le **note**, dove **<note>** può avere, oltre a **type** ed **n**, **resp** per indicare il responsabile (*autore, curatore,...*), **place** per indicare in che posizione appariva nel testo originale (*intralinear, interlinear, sinistra, destra, piede pagina, fine*), **target** per il punto di inserimento della nota e **targetend** per il punto di fine, **anchored** se il testo originale mostra l'esatto punto di riferimento per nota. Dubbi? Normale. Cerca ogni volta sulle guideline TEI: gli esempi servono a concretizzare nozioni altrimenti tutte uguali.

Gli **interventi editoriali** vanno messi in:

<corr> se si tratta di una correzione, con **sic** per l'errore originale nel testo, **resp** per il responsabile della correzione e **cert** per il grado della certezza della correzione;

<sic> se è un frammento di testo errato o impreciso, con **resp** e **cert** (come sopra) ma con **corr** ad indicare

la correzione dell'errore palese nel testo.

Se ci interessa studiare la normalizzazione del testo, le nozioni di cui sopra possono essere supportate da **<orig>** per la forma originale con **reg** per la regolarizzazione e **resp**, o al contrario **<reg>** per una regolarizzazione con **orig** e **resp**.

Ho notato che il testo ha o ha avuto (in revisioni o edizioni diverse) **parti mancanti**:

<add> contiene elementi inseriti da qualcuno, viene affiancato da **place** (*intralinea, sopralinea, sopra, sotto, ma anche insieme come interlinea sinistra...*);

<gap> indica una omissione del testo per motivi descritti di solito nella teiHeader (criteri di selezione) o perché illeggibile o parziale, viene affiancato da **desc** (si riferisce alla parte omessa) e **resp**;

**** invece descrive una parte cancellata nell'opera originale da parte dell'autore/revisore/copista/commentatore, viene affiancato da **type**, **status** (in caso di cancellature errate) e **hand** (riferito a chi ha fatto la cancellatura);

<unclear> si riferisce a cose illeggibili o incomprensibili, dunque non codificabili, viene affiancato da **reason** (perché non posso riportare il brano?) e **resp** (chi ha dichiarato quella parte non codificabile).

Mi interessa la cronologia, come **successione degli eventi**: date, ore.

<date> con **calendar** (es. *gregoriano*) e **value** espresso in gg-mm-aaaa; **<time>** con **value** in hh:mm:ss.

In caso di **numeri** rilevanti: **<num>** con **type** (*frazione/ordinale/percentuale/cardinale*) e **value**.

La codifica delle **liste** avviene con **<list>**. I suoi **type** possono essere: *ordinata* per elementi numerati o in ordine alfabetico, *puntata* per elenchi puntati, *glossario* per termini tecnici marcati ognuno con il suo elemento **<label>** e **<item>** (quest'ultimo per descrivere), infine *semplice* per liste non marcate.

Siccome i singoli elementi si segnalano con **<item>**, è opportuno mettere prima degli item un elemento **<head>** che faccia da titolo alla lista. Quest'ultima parte è importante nella creazione del peritesto che riguarda la bibliografia. La numerazione, in base alla lista, avviene con **n** o raramente con **<label>**, l'importante è non confondere tra loro i tipi e gli stili di codifica delle liste.

Più complesse dalle liste sono le tabelle.

<table> con **rows** e **cols** in generale

<row> con **role** (informazione nella riga) più **<cell>** con **role** (descrizione della cella), **label** (descrizione dell'elemento) e **data** (valore del dato); **cols** indica il numero di colonne che la cella occupa e **rows** quello delle righe.

Immagini varie vanno codificate con **<figure>** con i suoi attributi vari (consultare il sito in base alla necessità) e **<figdesc>** per la descrizione.

Se devo interpretare o **analizzare il testo** mi possono essere utili le seguenti codifiche.

In campo linguistico:

<s> è unità testuale, con **type** (dichiarativa, interrogativa, ecc).

In campo interpretativo:

o usiamo **<seg>**, che indica un segmento dove il suo **type** può essere qualsiasi cosa noi gli affibbiamo, ma è dunque opportuno spiegare questi attributi nella teiHeader;

o usiamo **<interp>** per un'annotazione interpretativa riferita a intervalli di testo, associata a **type** (il tipo di fenomeno rilevato *es figura retorica, personaggio, allusione, tema...*), **value** (il fenomeno rilevato), **resp** (di chi interpreta) e **inst** (istanza del testo a cui si applica l'interpretazione).

Spero vi sia utile. Di certo l'importante è settare uno standard. La TEI è così libera da diventare personale, e la codifica dovrebbe essere al servizio della sua consultazione, non dello stile di chi codifica. Grazie a chiunque abbia avuto la pazienza di esercitarsi con me.