

La codifica è stata eseguita seguendo le raccomandazioni della *Text Encoding Initiative* (TEI), presentate nelle *Guidelines P5* (<https://tei-c.org/guidelines/p5/>). Le seguenti osservazioni non sostituiscono un'attenta lettura di questa guida, ma intendono presentare le scelte operate per la codifica del prototipo di *digital scholarly edition* della *Storia d'Italia* di Francesco Guicciardini, affinché possano servire da modello per la codifica del resto dell'opera.

Le trascrizioni che stanno alla base dell'edizione sono state prodotte da Paola Moreno (Université de Liège) e Pierre Jodogne (Académie Royale de Belgique). La codifica XML-TEI è stata eseguita da Élise Leclerc (Université Grenoble Alpes). La modellazione e la progettazione dell'edizione scientifica digitale sono state realizzate da Élise Leclerc e Samantha Saïdi (Triangle UMR 5206, ENS de Lyon) in collaborazione con Paola Moreno, Hélène Miesse (Université de Liège) e i commentatori del gruppo di ricerca. Le visualizzazioni sono state sviluppate a partire dalla versione 1.1 dello strumento *Edition Visualization Technology* (EVT), progettato e sviluppato sotto la direzione di Roberto Rosselli del Turco. Lo strumento è stato adattato alle esigenze del progetto da Samantha Saïdi, che ha anche adattato gli scenari di trasformazione XSLT e i fogli di stile CSS.

Promemoria dei principi di base della codifica XML-TEI:

1. Gli elementi che danno un significato ai caratteri sono delimitati da marcatori di apertura e di chiusura, identificabili dalle parentesi uncinata (per esempio `<add>testo aggiunto</add>`). Ad ogni marcatore di apertura deve corrispondere un marcatore di chiusura. In alcuni casi, un singolo marcatore (detto marcatore autochiudente) basterà a fornire le informazioni. Per esempio `<pb/>` (salto di pagina) significa “qui si cambia pagina”, `<lb/>` (salto di riga) significa “qui si cambia riga”.
2. I marcatori contenenti caratteri seguono un rigoroso principio di incastramento, come delle matryoska: `<add>Qui inizia un'aggiuntaquesta parte è stata cancellata mentre l'aggiunta continua fino a questo punto.</add>`. Una sintassi che presentasse delle sovrapposizioni tra i marcatori non sarebbe valida (per esempio `*<add>Qui inizia un'aggiuntaquesta parte è stata cancellata e l'aggiunta termina qui </add>e la cancellazione qui `).
3. Il descrittore principale tra le parentesi uncinata di un marcatore di apertura e quello di chiusura è detto *elemento*. In “`<del rend="underline">sequenza di caratteri`”, l'elemento è `del`: delimita e qualifica la sequenza di caratteri che i marcatori contengono – i quali di per sé non hanno un significato per la macchina – come una porzione di testo che è stata cancellata. L'elemento viene usato per fornire informazioni sullo stato (qualità, posizione, ecc.) dei caratteri che contiene. Se il lettore è un programma informatico, il codice deve includere istruzioni specifiche per trattare in conseguenza il contenuto testuale del marcatore. I nomi, le proprietà e gli usi degli elementi sono censiti e descritti nella guida TEI, per facilitare un loro uso standardizzato, da un'edizione digitale all'altra.
4. Gli elementi possono avere *attributi* che forniscono informazioni supplementari. Pure gli attributi sono censiti e descritti nella guida TEI. Consentono di distinguere tra `<del rend="overstrike">testo soppresso cancellando i caratteri` e `<del rend="underline">testo soppresso sottolineando i caratteri`. In questo esempio, l'elemento `` fornisce informazioni sul significato dell'operazione, mentre il valore (“underline” o “overstrike”) dell'attributo `@rend` specifica come l'operazione sia stata materializzata sul foglio. Queste informazioni possono essere elaborate in maniera diversa nelle visualizzazioni dell'edizione digitale, a seconda delle scelte editoriali.

1. Struttura generale del documento XML-TEI

Il file XML-TEI possiede una struttura specifica, comune ai documenti di questo tipo: si apre con elementi dichiarativi (lingua, caratteri, schema associato...).

```
<?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE TEI SYSTEM "tei_all.dtd">
<TEI xml:id="SI" xmlns="http://www.tei-c.org/ns/1.0"> [... ] </TEI>
```

È poi strutturato in diverse parti. Il `<teiHeader>`, obbligatorio, contiene una serie di metadati sul file, le fonti, la codifica e le revisioni. Ad ogni nuova fase di intervento sul file, i metadati vanno completati (nomi degli autori degli interventi, date delle revisioni,...). Una sezione opzionale `<facsimile>`, suddivisa a sua volta in tante `<surface>` quante sono le riproduzioni fotografiche dei manoscritti inclusi nell'edizione, identifica le immagini utilizzate e permette di stabilire dei collegamenti tra il testo dell'edizione e queste immagini nello strumento di visualizzazione, secondo il modello seguente:

```
<facsimile>
<surface xml:id="surf_Ms05b_folio_1r" corresp="#SI_Ms05b_folio_1r">
  <graphic height="1800px" url="immagini\SI_Ms05_folio_1r.jpg" width="1237px"/>
  <zone corresp="#SI_Ms05b_folio_1r_01" lrx="1052" lry="211" rend="visible" rendition="Line" ulx="261"
uly="156" xml:id="SI_Ms05b_line_1r_01"/>
</surface></facsimile>
```

Segue la sezione obbligatoria `<text>`, composta da un corpo principale di testo (`<body>`) diviso in tante parti (`<div>`) quanti sono i manoscritti rappresentati (otto, in questo caso), più l'edizione di riferimento curata da Gherardi (1919). La struttura del file XML-TEI incontra allora la rappresentazione della struttura logica e fisica della fonte rappresentata nell'edizione.

```
<text>
<body>
  <div n="Ms01_AGF_I1b" subtype="edition_text" type="verse" xml:id="SI_Ms01">
  [...] </div></body></text>
```

2. Struttura logica e fisica del documento.

Ogni divisione (`<div>`) corrisponde a un'unità codicologica. Ha un identificatore unico (`xml:id="SI_Ms01"`) e un numero (`n="Ms01_AGF_I1b"`). Il valore "verse" è stato dato all'attributo `@type` (invece di "prose") perché in EVT la configurazione per i versi era più adatta al progetto (*display* riga per riga, numerazione delle righe, ecc.). Ogni divisione contiene un'entità `<p>` che corrisponde a un blocco "logico" di testo (di solito un paragrafo).

```
<div n="Ms01_AGF_I1b" subtype="edition_text" type="verse" xml:id="SI_Ms01">
  <p><pb n="ms01_1r" xml:id="SI_Ms01_folio_1r"/>
  <lb facs="#SI_Ms01_folio_1r_01" n="1" xml:id="SI_Ms01_1r_lb01"/>
  <milestone unit="section" n="04" xml:id="Ms01_04"/>
  [texte du segment 4] </p> </div>
```

All'interno di questo insieme viene codificata la struttura fisica del documento: la divisione in fogli e pagine è codificata usando `<pb/>`. Ogni pagina ha un attributo che le dà un identificatore univoco (`@xml:id`) e un numero (`n="ms01_1r"`), il cui valore appare nei menù dello strumento di visualizzazione. Ogni pagina è poi suddivisa in righe per mezzo di `<lb/>`. Ogni riga ha anch'essa un identificatore univoco, un numero e un attributo (`@facs`) che ne collega il contenuto all'immagine corrispondente nella sezione `<facsimile/>`. È stata poi aggiunta un'ulteriore suddivisione logica, con `<milestone/>`, per delimitare i segmenti di testo (dotati pure loro di un identificatore unico), stabiliti a partire dall'edizione di riferimento (Gherardi 1919), al fine di mettere a confronto nella sezione "Segmenti" del sito i diversi stati dello stesso brano.

II. RAPPRESENTAZIONE DELLA FONTE PRIMARIA

1. Rappresentazione fisica del documento.

La visualizzazione semi-diplomatica dell'edizione mira a riflettere il contenuto e l'impostazione del testo sul foglio, proponendo una trascrizione dei manoscritti (molto più leggibile della grafia tormentata dell'autore) che rispetti la disposizione dei caratteri e la loro forma. Questa visualizzazione si fonda sulle informazioni codificate nel file XML-TEI utilizzando una serie di marcatori che descrivono la natura, la posizione e/o la forma dei caratteri presenti nel foglio.

L'elemento `` (soppresso) è usato per indicare che un brano è stato cancellato durante il processo di scrittura. Quest'operazione può essere eseguita con diversi segnalatori mentre l'informazione viene codificata grazie ai valori dell'attributo `@rend` ("underline" o "overstrike"). Se la soppressione è stata immediata, il brano è inquadrato da `<del instant="true">`.

L'elemento `<hi>` (evidenziato), combinato con l'attributo `@rend`, è usato per descrivere brani il cui aspetto (ad es. sottolineato o centrato) è diverso dal testo circostante, senza però che quest'aspetto sia dovuto ad un'operazione codificata da altri marcatori.

L'elemento `<add>` (aggiunta) è usato per indicare che un brano è stato aggiunto durante il processo di scrittura. La posizione di tale aggiunta sul foglio è descritta con l'attributo `@place`, che può avere i seguenti valori: "above" (è nell'interlinea), "margin" (è nel margine, che è di default quello sinistro), "top" (è nel margine superiore), "bottom" (è nel margine inferiore), "superimposed" (una lettera è stata cambiata in un'altra lettera), "inline" (quando l'aggiunta è stata fatta direttamente sulla riga). Gli attributi `@n` e `@xml:id` sono stati talvolta utilizzati per assegnare un numero e/o un identificatore alle aggiunte: il numero indica l'ordine di apparizione in un'area del foglio (margine superiore o inferiore), mentre l'identificatore può essere utile per identificare operazioni congiunte durante il processo di codifica (tale porzione aggiunta nel margine superiore è una variante di tale porzione eliminata nel corpo del testo, che è stata poi oggetto di un'altra variante nel margine superiore...). L'attributo `@seq` (sequenza, con un valore numerico) è stato talvolta usato per codificare l'ordine semantico delle aggiunte marginali, sia quando ci sono aggiunte in più margini, sia quando le aggiunte in un margine non seguono l'ordine abituale (dall'alto verso il basso).

L'elemento `<metamark>` è usato per indicare che il carattere o i caratteri che contiene sono segni diacritici, aggiunti dall'autore per indicare come il testo debba essere letto, per identificare le correzioni che fa o per indicare brani alternativi. L'attributo `@funzione` è usato per distinguere i diversi casi: o è un segno che indica un'operazione di "insertion" (il segno rimanda a un altro segno identico in uno dei margini, seguito dal testo sostitutivo, o indica l'ordine di due porzioni da invertire) o un'"alternative" (`<metamark funzione="alternative"></metamark>` incornicia il segno

diacritico *vel*, e indica che quanto segue è una possibile alternativa). L'attributo `@n` può essere usato per identificare le corrispondenze tra i segni. In un'occasione, Guicciardini fa un commento in margine. Per maggiore comodità nella trasformazione dei dati (essa doveva apparire solo nella versione semi-diplomatica), abbiamo usato la seguente codifica: `<metamark function="commentary" resp="#FG">commento</metamark>` invece di un elemento specifico, come `<note>`.

L'esempio *infra* illustra gran parte dei marcatori e attributi evocati finora, eppure `<subst>` (sostituzione). Serve a descrivere un'operazione di correzione nel suo complesso, combinando una soppressione `` e un'aggiunta `<add>`. I due elementi sono obbligatori nella sostituzione (`<subst>testo cancellato<add>testo aggiunto</add></subst>`) ed è possibile incastrare più sostituzioni:

```
<del xml:id="change0523">
  <add place="bottom" n="1"><metamark function="insertion" n="Met0502"> .A.
</metamark>Sarà oltre a questo <rs type="concept" ref="cognitione">cognitione</rs> piena di
utilità perché ciaschuno potrà prenderne <lb/><del>et</del> per se proprio et per beneficio <rs
type="concept" ref="pubblico">pubblico</rs> molti salutiferi documenti<pc> : </pc> donde <del
rend="overstrike">chiarissim</del> per <lb/>innumerabili exempli chiarissimamente
apparirà</add></del>
<add place="bottom" n="2" xml:id="change0524"><metamark function="insertion"
n="Met0502">A.</metamark><del rend="overstrike">Potrà</del> d<add
place="superimposed">D</add>alla <rs type="concept" ref="cognitione">cognitione</rs>
d<subst><del rend="overstrike">i</del><add place="superimposed">e'</add></subst>
<add place="above">quali</add> casi tanto varii et tanto gravi <add place="above">potrà</add><del
rend="overstrike">prendere</del> ciaschuno et per se <lb/>proprio et
  <subst><del rend="overstrike">per beneficio <rs type="concept" ref="pubblico">
pubblico</rs></del><add><metamark function="alternative" rend="ital">vel</metamark>
per bene <subst><del rend="overstrike">di altri</del><add place="above"><rs
type="concept" ref="pubblico">pubblico</rs></add></subst></add></subst>
<del rend="overstrike">prendere</del><add place="above">prendere</add> molti salutiferi <rs
type="concept" ref="documenti">documenti</rs><pc> : </pc>
  <subst><del rend="overstrike">donde</del><add place="above">onde </add> </subst>
<lb/>per innumerabili <subst><del instant="true" rend="overstrike">ex</del><add
place="above">exempli</add></subst>
  <subst><del rend="overstrike">evidentemente <metamark function="alternative"
rend="ital">vel</metamark> chiarissimamente</del><add place="above">
evidentemente </add></subst>apparirà </add>
```

Tali operazioni di sostituzione sono fondamentali per l'edizione digitale della *Storia d'Italia*, perché le tre rappresentazioni proposte nella sezione "Manoscritti" del sito dipendono dalla loro sequenza. La rappresentazione "semi-diplomatica" piglia tutti gli elementi presenti nel documento (cancellazioni e aggiunte) e li distribuisce nelle aree dello schermo che corrispondono alle aree del foglio, mentre la rappresentazione "primo getto" piglia quel che corrisponde al primo stato del testo di ogni foglio, il quale si trova all'inizio delle catene di sostituzioni (la prima ``, o le parti di testo invariate), mentre la rappresentazione "testo revisionato" piglia le parti invariate e le informazioni contenute nell'ultima `<add>` della serie.

Per i pochi casi d'interversione di due porzioni di testo, per ottenere una maggiore semplicità di elaborazione, si è preferito l'elemento `<seg>` (segmento) con valori "orig" o "reg" per l'attributo `@type` agli elementi `<listTranspose>` e `<transpose>`. L'ordine iniziale è codificato con

@type="orig" e appare nelle rappresentazioni “semi-diplomatica” e “primo getto”, mentre il risultato dell’interversione è codificato con @type="reg", e appare solo nel "testo revisionato".

```
<lb facs="#SI_Ms04b_folio_2v_02" n="2" xml:id="SI_Ms04b_lb_2v_02"/>
<seg type="orig"><rs type="concept" ref="lacerare">lacerata</rs></seg>
<metamark function="insertion" place="above" n="Met0401"> b </metamark> <seg type="orig">si
miserabilmente</seg>
<metamark function="insertion" place="above" n="Met0402"> a </metamark>
<seg type="reg"> si miserabilmente lacerata </seg>
```

Sul sito, le sostituzioni sono evidenziate; quando l’operazione di sostituzione riguarda sia una porzione del corpo del testo che caratteri marginali, la prima è stata evidenziata mediante <hi rend="grey" resp="editor">porzione evidenziata in grigio</hi>) per rendere l’operazione più visibile e leggibile. È un esempio d’intervento non autoriale ma editoriale.

2. Interventi editoriali.

Le poche abbreviazioni usate da Guicciardini sono state sciolte in tutte le trascrizioni, compresa quella semi-diplomatica, e quindi non sono state codificate in maniera specifica.

Le lettere maiuscole e la punteggiatura (segnalata dall’elemento <pc>) sono state modernizzate nelle rappresentazioni “primo getto” e “testo revisionato”, usando gli elementi <choice>, <orig> e <reg>. L’elemento <choice> consente di raggruppare due forme alternative: una originale, inclusa in <orig>, e una forma normalizzata, in <reg>.

```
<choice><orig><pc> / </pc></orig><reg><pc>, </pc></reg></choice>
<lb facs="#SI_Ms04b_folio_2v_06" n="6" xml:id="SI_Ms04b_lb_2v_06"/>lo <rs type="org" ref="Impero
romano">Imperio</rs><choice><orig>R</orig><reg>r</reg></choice>omano
```

Per via dello stato di conservazione dei manoscritti, la lettura di alcune parole è a volte incerta: lo si segnala con <unclear>, e viene rappresentato sul sito da parentesi “testo certo(testo incerto)”. L’attributo @reason è usato per indicare se sia dovuto alla grafia (“illegible”) o all’inchostro sbiadito (“faded”). Se i caratteri sono totalmente illeggibili, si usa l’elemento <gap>, rappresentato sul sito da “[...]”. In caso di omissione volontaria da parte dell’autore, ad esempio uno spazio lasciato vuoto per essere completato più tardi, è stato usato <space>, rappresentato sul sito da “[***]”. Qualora una parte del testo sia stata congetturata dagli editori (nei casi di danni materiali, ad esempio), è stato usato l’elemento <supplied>, rappresentato sul sito da “testo<congettura>”.

L’elemento <note> è stato usato per aggiungere informazioni sulla modifica: i valori "gloss" o "edit" dati all’attributo @type, combinati con due richiami di nota (@n="o" o @n="*"), consentono di distinguere tra note sul contenuto del testo (ad. es. segmenti che non esistono ancora in una redazione) e quelle che riguardano la forma del testo. Esse consentono di superare i limiti della rappresentazione mimetica del foglio (ad esempio, per gli interventi di correzione per sottolineatura che sono poi annullati cancellando il tratto di penna). L’attributo @xml:lang è usato per distinguere tra lingue diverse.

```
<note type="gloss" resp="#EL" n="o" xml:lang="fr">[Segments précédents inexistant dans cette rédaction] </note>
<note type="edit" xml:lang="fr" n="*">[La correction, indiquée par le soulignement, a ensuite été annulée en barrant le
trait]</note>
```

In previsione di link verso i commenti e della creazione di vari indici, l'elemento `<rs>` (catena di riferimento) è stato utilizzato per codificare i concetti, i nomi di persone, luoghi ritenuti degni di interesse dal *team* di commentatori. I valori dell'attributo `@type` permettono di distinguere tra i diversi casi ("`person`", "`place`", "`concept`"), mentre il valore dell'attributo `@ref` contiene una versione di riferimento per il contenuto del marcatore che permette, ad esempio, di ovviare alle varianti ortografiche, ecc.

```
<rs type="person" ref="Innocenzio VIII">Innocentio 8° Pontefice Romano</rs>
<rs ref="Innocenzio VIII" type="person">Innocentio Octavo Pontefice Romano</rs>
```

L'elemento `<date>` è stato usato per codificare le date, in previsione di futuri indici e dello sfruttamento di tali informazioni da parte dei commentatori. L'attributo `@when` consente di specificare un valore normalizzato per la data, indipendentemente dalla sua grafia.

```
<lb facs="#SI_Ms03a_folio_1r_15" n="15" xml:id="SI_Ms03a_lb_1r_15"/>l'anno della salute cristiana <date
when="1490"><choice><orig>M</orig><reg>m</reg></choice>ille <choice> <orig>Q</orig> <reg>q</reg>
</choice>uattrocento novanta</date>
<date when="1490">mille quattrocento novanta</date>
```

3. Prospettive.

Rendere conto delle diverse mani è una delle direzioni di lavoro prioritarie. Per ora, sono stati codificati soltanto gli interventi di Francesco Guicciardini ("`#FG`") e del nipote Agnolo Guicciardini ("`#AG`") nel manoscritto Palatino 166, grazie all'attributo `@resp` dell'elemento `<add>`. Queste informazioni non si vedono ancora sul sito, poiché si aspetta di aver identificato meglio tutte le mani (comprese quelle dei segretari) e di aver trovato una soluzione soddisfacente per renderle visibili nell'edizione.

Un finanziamento *ad hoc* permetterebbe poi di ottenere riproduzioni fotografiche dei manoscritti conformi agli standard internazionali in questo campo, in particolare per il Palatino 166, il che consentirebbe di sfruttare appieno le potenzialità dello strumento di visualizzazione dei facsimili.

Nel frattempo, ci si augura che il lavoro svolto sull'esordio, parte particolarmente complessa dell'opera (per via della grande varietà d'interventi autoriali presenti nelle sue numerose redazioni) abbia permesso di sviluppare un modello di codifica facile da applicare all'insieme del capolavoro del Guicciardini.