
TEI markup for digital facsimiles and diplomatic editions

Roberto Rosselli Del Turco
Dipartimento di Studi Umanistici
Università di Torino
roberto.rossellidelturco@unito.it





SUMMARY

TEI modules for Digital philology

The Manuscript Description module

The parallel transcription

The digital facsimile

The embedded transcription

Character encoding and the TEI gaiji module

Edition levels



TEI modules for Digital philology



A short reminder

- **diplomatic** edition: transcription of the text of a witness respecting the original layout and spelling, without any kind of correction (evident errors) or other editorial interventions (expansion of abbreviations) → “textual facsimile”
- **diplomatic-interpretive** edition: still respecting the original text, the most obvious errors are corrected, certain spelling peculiarities are regularized (word subdivisions), abbreviations are expanded, etc. → documentary eds.
- **critical** edition: after the collation of all witness transcriptions, the *stemma codicum* is established and an attempt is made to reconstruct the original text by comparing the variants of the most valid witnesses → Lachmann’s method

Custom TEI schema for digital philology

- the required elements are in part already available in the basic TEI modules (core), but for textual criticism are necessary the following:
 - ***msdescription*** manuscript description
 - ***trans*** transcription of primary sources
 - ***textcrit*** critical apparatus
 - ***analysis*** elements for text analysis
 - ***gaiji*** non-standard characters
- depending on the type of markup also useful: *linking, certainty, figures, namesdates, verse* → training exercise with **Rome**

Why create a custom TEI schema

- you could use tei_all to have all TEI elements available all the time, or use Oxygen XML which has internal support for TEI (no schema association needed)
- creating your own schema is the smarter move because
 - a smaller schema means **faster validation**, greater efficiency
 - smart editors will present a **shorter list** of available elements: faster to choose the right one, less error-prone
 - easier to have the schema fit the needs of an **annotation model**
 - learning how to create a schema also means being able to modify and **further customize** it should the need arise



The Manuscript Description module



Manuscript description module

- adding this module makes available an **<msDesc>** element that can be used almost anywhere, but has its natural place in **<sourceDesc>**
- **<msDesc>** has a potentially complex structure, with many dedicated sub-elements
- can be used for a quick description, including for cataloging purposes, of a manuscript, or to include a proper paleographic and codicological introduction
- more information in the *Manuscript Description* chapter of the *Guidelines*: <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/MS.html>

Fundamental <msDesc> sub-elements

- <**msIdentifier**> information required to identify the manuscript
- <**msContents**> describes the intellectual content of a manuscript or manuscript part as free text (<p> elements) or as a series of structured manuscript items (<**msItem**> elements)
- <**physDesc**> full physical description of a manuscript or manuscript part, optionally based on more specialized elements
- <**history**> elements describing the full history of a manuscript
- <**additional**> additional information about the manuscript
- <**msPart**> information about an originally distinct manuscript

```
<sourceDesc>
<msDesc>

    <!-- Informazioni sulla collocazione fisica del Codice -->
    <msIdentifier>
        <settlement>Sarzana</settlement>
        <repository>Archivio del Capitolo della cattedrale</repository>
        <msName>Codice Pelavicino</msName>
    </msIdentifier>

    <!-- Informazioni su autore e lingua -->
    <msContents>
        <msItem>
            <textLang mainLang="latin">Latino</textLang>
            <colophon>Il codice si compone di 426 cc. numerate e 20 cc. non numerate;
                contiene diversi testi tra cui il <emph>Liber Iurium</emph> della Chiesa di
                <placeName>Luni</placeName>. Principale responsabile della sua redazione
                fu il vescovo di Luni <persName>Enrico da Fucecchio</persName>, che salì al
                soglio episcopale nel 1273 e vi rimase fino al 1289. Al fine di
                salvaguardare i beni e i diritti della Chiesa di
                <placeName>Luni</placeName>, il vescovo <persName>Enrico</persName>
                riorganizzò infatti gli uffici della curia, fece compilare un inventario
                generale dell'archivio ecclesiastico e attivò uno scrittorio nel quale
                lavorarono diversi amanuensi, da cui uscì appunto buona parte del codice. Il
                nome di <emph>Codice Pelavicino</emph> deriva da una piccola parte del
                codice stesso, il cosiddetto <emph>Liber Magister</emph>, cioè il libro dei
                redditi e dei proventi della Chiesa di <placeName>Luni</placeName>, che è
                riportato nel codice nella copia tardoduecentesca, ma che originariamente
                venne compilato al tempo di <persName>Uberto Pelavicino</persName>, vicario
                in Lunigiana dell'imperatore <persName>Federico II</persName> nel quarto
                decennio del secolo XIII. Il nome tuttavia si impose presto nell'uso per
                indicare l'intero codice, compreso il <emph>Liber Iurium</emph>, come
                testimonia un documento del 1487 trascritto nello stesso
                manoscritto.</colophon>
        </msItem>
    </msContents>
```

<!-- Descrizione fisica del codice.-->

```
<physDesc>
  <objectDesc>
    <supportDesc>
      <p>Il volume è rilegato in <material>pelle</material> con piatti in legno (<width>360</width> * <height>255</height> * <depth>5mm</depth>); il fondello presenta cinque nervature; la chiusura è data da una correggia con fibbia in metallo. La guardia anteriore è formata da due fogli di carta relativi al restauro del 1984, operato presso l'<emph>Istituto centrale per la patologia del libro "Alfonso Gallo"</emph> di Roma (ACS <emph>Corrispondenza 1982-1987</emph>, n.6). Seguono 5 fogli di carta (10 cc.) con una fine vergellatura e filigrana a mezzaluna databile alla seconda metà XVI-inizio XVII secolo. La risguardia è di un foglio di carta di epoca recente incollato sulla coperta.</p>
      <p>La guardia posteriore è formata da due fogli di carta recente e la risguardia da un foglio di carta recente incollato alla coperta. Precedono tre fogli di carta identici a quelli della guardia anteriore (6 cc.) recanti la medesima filigrana, preceduti da un foglio irregolare di pergamena con un crittogramma del secolo XVI numerato a matita 426. I fogli cartacei del XVI-XVII secolo misurano <width>475</width> * <height>335 mm</height>.</p>
      <p>Il codice conta 37 fascicoli. Molti fascicoli sono numerati con numeri romani sia cardinali che ordinali.</p>
      <p>Il 1° fascicolo è privo di indicazioni che ne stabiliscano la posizione nel codice; il 2° non presenta indicazioni, ma originariamente possedeva probabilmente la numerazione <hi rend="bold">I(us)</hi> scomparsa in seguito alla rifilatura; il 3° e il 4° presentano nel margine inferiore, nel verso dell'ultima carta, <hi rend="bold">II(us)</hi> e <hi rend="bold">III(us)</hi> in inchiosto nero; il 5° fascicolo non presenta indicazioni, ma possedeva probabilmente la numerazione <hi rend="bold">IIII(us)</hi>, scomparsa in seguito alla rifilatura.</p>
      <p>La numerazione dei 6° al 20° è in inchiostro rosso, a cui segue dal 21 in poi una numerazione in inchiostro nero. Il 21° fascicolo (XVI) presenta una nota, probabilmente della medesima mano che ha numerato i fascicoli del liber iurium, in cui si avverte di una lacuna: <emph>hic deficiunt
```

```
</supportDesc>
</objectDesc>
<handDesc hands="6">
    <p>Nel manoscritto sono riconoscibili diversi scrittori. Si elencano qui i
        principali fino ad ora (ottobre 2015) riconosciuti. L'elenco definitivo
        verrà pubblicato al termine del lavoro di trascrizione ed edizione
        critica.</p>
    <p><hi rend="bold">(1)</hi> Un amanuense di cui ignoriamo il nome ha iniziato a
        scrivere il 12 dicembre 1287 e ha redatto i primi 22 documenti del
            <emph>Liber Iurium</emph> (nn. 27-48, cc. 57r.-70r).., dopodiché il lavoro
        fu interrotto; la scrittura di questo primo amanuense è riconducibile al
            tipo della <emph>rotunda italiana</emph>, avvicinabile alla <emph>littera
            bononiensis</emph>, propria dei manoscritti universitari bolognesi, ma
        diffusa anche in molti altri centri italiani; è più tondeggiante, la
        caratteristica del tratto spezzato, tipica della gotica, è poco avvertibile,
        mentre le lettere sono più addossate l'una all'altra. Le principali
        differenze rispetto alla scrittura n° 2 (di <emph>magister</emph> Egidio)
        sono così riassumibili:</p>
    <p>
        <list>
            <item>la <emph>g</emph> rotonda, assomigliante ad un 8, l'occhiello
                inferiore è rotondo anziché schiacciato a formare un triangolo;</item>
            <item>la <emph>p</emph> con il significato di <emph>per</emph> reca il
                taglio orizzontale alla base e non a metà dell'asta;</item>
            <item>il segno tachigrafico per <emph>et</emph> è sempre a forma di 7, ma
                non reca il taglio orizzontale; non compare mai il segno a forma di 9
                con significato di <emph>cum/con</emph>;</item>
            <item>le <emph>i</emph> non presentano apici;</item>
            <item>la <emph>y</emph> è priva di puntino al centro della
                biforcazione;</item>
            <item>accanto alla <emph>d</emph> onciale, con asta abbattuta e quasi
                orizzontale, lo scriba adopera sovente la <emph>d</emph> diritta
                pressoché assente in Egidio.</item>
        </list>
    </p>
```

```
<!-- Informazioni storiche sul codice -->
<history>
    <summary>
        <p>Si compone di 426 carte e contiene in copia documenti dall'anno
            <date>900</date> (diploma di Berengario re d'Italia a favore del vescovo
            lunense) fino al <date>1289</date>.</p>
    </summary>
    <origin>
        <p>La maggior parte del manoscritto risale alla seconda metà <origDate
            notBefore="1199" notAfter="1299">XIII secolo</origDate>.</p>
        <p>Le diverse parti in cui ora è composto il codice vennero assemblate in una
            prima rilegatura plausibilmente conservata nella sede del vescovo di Luni,
            ossia nel <origPlace><emph>castrum Sarzane</emph></origPlace>, attuale
            Sarzanello, prima del 1487.</p>
        <p>Come ci informa un documento trascritto nello stesso codice (CP cc. 54v-56v,
            n. 26), a quell'epoca il vescovo di Luni risiedeva temporaneamente a
            Pontremoli, presso la chiesa di S. Colombano, a causa della guerra di
            Sarzana, 49-52). Da qui il vescovo Tommaso aveva indirizzato una supplica a
            papa Innocenzo VIII affinché attribuisse ai documenti trascritti nel codice
            valore di prova nelle contestazioni giudiziarie. Dietro mandato papale il
            giureconsulto Tomaso Amadeo di Ferrara, in Modena, autenticò il
            <emph>liber</emph>, già detto all'epoca <emph>Pelavicinus</emph>, il 9
            agosto 1487. In quest'atto il codice è detto composto di 408 carte, che
            iniziano con l'indice (attuale c. 1, fascicolo 1) e terminano con un
            documento del 1289 (attuale c. 408v., fascicolo 35): questo significa che
            già nel 1487 mancavano dal volume i fascicoli dal XVII al XX. In un momento
            successivo, non noto, furono recuperati alcuni fogli di un fascicolo
            perduto, quasi certamente il XVII (attuali cc. 409r-414v. del fascicolo 36)
            e quindi inseriti in una nuova rilegatura. In questa nuova composizione al
            codice vennero probabilmente aggiunte le carte perse che ora formano il
            fascicolo 37 contenenti una redazione del <emph>Liber Magister</emph>
            differente da quella copiata da maestro Egidio.</p>
        <p>A una terza rilegatura dovrebbero appartenere i fogli di carta della guardia
            anteriore e posteriore dotati di una filigrana databile a cavallo tra i
            secoli XVI e XVII.</p>
    </origin>
</history>
```



The parallel transcription



Page layout

- a few general elements needed for a transcription
- <pb/> page begins: beginning of a new folio / page

<**pb** n="214v" xml:id="fol_214v"/>

- <lb/> line begins: beginning of a new text line

<l n="1"><**lb** n="7" xml:id="VB_lb_104v_07"/><hi
rend="init3.1"><g ref="#Hunc"/></hi>

- <cb/> column begins: beginning of a new column

<**cb** n="1"/><lb n="1" xml:id="VB_lb_104v_07"/>

Added / Deleted text, lacunae

- **<add>** one or more words/characters added in the text

this word is `<add place="supralinear">being</add>` added at a later time

- **** one or more words/characters deleted in the text

this one was `<del rend="overstrike">was` in excess and I deleted it

- **<gap/>** part of the text omitted, missing, or illegible

this `<gap reason="illegible" quantity="4" unit="chars"/>` is illegible

Damaged and/or unclear text, supplied text

- **<damage>** damaged text in the original document

this word has **<damage agent="water">discolored</damage>**
because of water damage

- **<unclear>** part of text interpretable with difficulty (mod. transcr)

<unclear reason="faded">this</unclear> can still be read
but with difficulty

- **<supplied>** text inserted by the editor because it was illegible in the
original or absent (part of transcr mod)

here **<supplied>something</supplied>** was missing

Combining transcription elements

```
<w>ge<lb n="24"/>bund<g ref="#eenl"/>nne</w>
<damage type="faded"><g ref="#inodotenl"/>n</damage>
<unclear><g ref="#sins"/>umne</unclear>
<gap reason="illegible" extent="unknown"/>
<supplied resp="scragg">cafertun
    þa eode <name type="person">Petrus</name>
    ond þes ilca</supplied>
```

From the Vercelli Book manuscript: Old English text, late West Saxon dialect, end of X century

Substitutions

- when a word is not simply deleted, but is also replaced by another one, you can use the **<subst>** element to explicit the sequence deletion → new text (see **<choice>**)

this word was **<subst>**

```
<del rend="overstrike">deleted</del>
<add place="supralinear">added</add>
```

<subst> at a later time

- also at the level of a single letter

```
<subst><del rend="overtype">t</del><add>T</add></subst>
```

Suggested values for attributes

- <add> **place** inline, supralinear, margin, etc.
hand author, scribe1, scribe2
- **rend** overstrike, subpunct, overtype, dotted
hand author, scribe1, scribe2
- <gap/> **reason** illegible, cancelled, missing, omissis, censored
hand author, editor
agent water, smoke, hole, missing page
unit chars, words, cm
- <unclear> **reason** parterased, ink blot
hand author, scribe1, scribe2
agent water, smoke, ink

Couples of editorial elements

- besides <supplied>, TEI offers a few editorial elements which are perfectly specular:
 - <sic> - <corr>
 - <orig> - <reg>
 - <abbr> - <expan>
- you can use them individually, couples are not mandatory
- but if you use both of them it is very much recommended to also use the <choice> element to make the connection explicit

Corrections and regularizations

- **<sic>** word or phrase believed to be incorrect, but reported "as is"

this word is <sic>eevidntly</sic> wrong

- **<corr>** correction of an incorrect word or phrase

this word has been <corr>corrected</corr> at a later time

- **<orig>** word or phrase deemed "nonstandard"

just tell me <orig>wr</orig> are <orig>u</orig> going

- **<reg>** normalized (regularised) word or phrase.

just tell me <reg>where</reg> are <reg>you</reg> going

Abbreviations and expansions

- **<abbr>** an abbreviated word, a brevigraph

ask <abbr>dott.</abbr> Rossi

in nomine Patris <abbr>7</abbr> Filii <abbr>7</abbr>
Spiritus Sancti

- **<expan>** expansion of an abbreviated form

ask <expan>dottor</expan> Rossi

in nomine Patris <expan>et</expan> Filii
<expan>et</expan> Spiritus Sancti

Abbreviations and expansions

- **<abbr>** use the **type** attribute to specify the abbreviation type

ask <abbr type="suspension">dott.</abbr> Rossi
in nomine Patris <abbr type="brevigraph">□</abbr>
Filiī <abbr type="brevigraph">□</abbr> Spiritus

- adding the transcr module more elements become available:
 - **<am>** letters with abbreviation signs
 - **<ex>** letters added when expanding an abbreviation

```
<abbr>sanct<am>ū</am></abbr>  
<expan>sanctu<ex>m</ex></expan>
```

Using <choice>

- <sic> - <corr>

this word is <choice>

<sic>eevidntly</sic>

<corr>evidently</corr>

</choice> wrong

- <orig> - <reg>

just tell me <choice><orig>wr</orig>

<reg>where</reg></choice> are you going

- <abbr> - <expan>

ask <choice><abbr>dott.</abbr>

<expan>dott<ex>or</ex></expan></choice> Rossi

Important attributes for editorial annotation

- for <add> and we have the **hand** attribute to specify, if necessary, the author of the change to the original text
- for some editorial interventions it may be important to specify the person responsible (especially for <corr> and <supplied>!) and the degree of certainty (to filter the text)
- to do this use the **resp** and **cert** attributes:

```
this word is <choice>
    <sic>shorely</sic>
    <corr resp="rrdt" cert="100">surely</corr>
</choice> a writing mistake
```

- the value of **resp** may defer to a <respStmt>:

```
<corr resp="#rrdt">surely</corr>
```

Other useful elements

- <gb/> gathering begins: indicates the beginning of a new gathering (booklet) in the transcript
- <space> used to mark the presence of significant space (e.g. space left by scribe to insert an illuminated initial) (NB: *transcription Representation of Primary Sources* module)
- <pc> punctuation character(s): contains one or more characters that constitute a form of punctuation in the text (NB: *analysis Simple Analytic Mechanisms* module)

Other useful elements

- **<restore>** to record the scribe's (or author's) second thoughts about a deletion or other changes
- **<secl>** text that is presumed to be original, but is not in its proper location

<secl>

```
<l n="15" xml:id="l15">Alphesiboea suos ulta est pro  
coniuge fratres,</l>  
<l n="16" xml:id="l16">sanguinis et cari vincula rupit  
amor.</l>
```

</secl>

- **<surplus>** text deemed superfluous or spurious

Other useful elements

- <**seg**> arbitrary segment: used to mark any portion of text, use **type** to specify semantic content (NB: *linking* *Linking, Segmentation and Alignment* module)
- <**w**> word: mark a word at the grammatical level, useful attributes: **lemma** to indicate the lemma and **lemmaref** to establish a link to a separate dictionary, including online ones (NB: *analysis Simple Analytic Mechanisms* module)
- <**c**> character: marks a single character in the text (NB: *analysis Simple Analytic Mechanisms* module)



Editto di Rotari (Codex Vercellensis CLXXXVIII)

```
<msDesc xml:id="CE34">
  <msIdentifier>
    <country>Italia</country>
    <settlement>Ivrea</settlement>
    <repository>Biblioteca Capitolare di Ivrea</repository>
    <idno>34</idno>
  </msIdentifier>

</msDesc>
<msDesc xml:id="CV188">
  <msIdentifier>
    <country>Italia</country>
    <settlement>Vercelli</settlement>
    <repository>Biblioteca e Archivio Capitolare di Vercelli</repository>
    <idno>CLXXXVIII</idno>
  </msIdentifier>
  <msContents>
    <textLang mainLang="lat">The primary language in this manuscript is Latin, with a few isolated Langobardic words.</textLang>

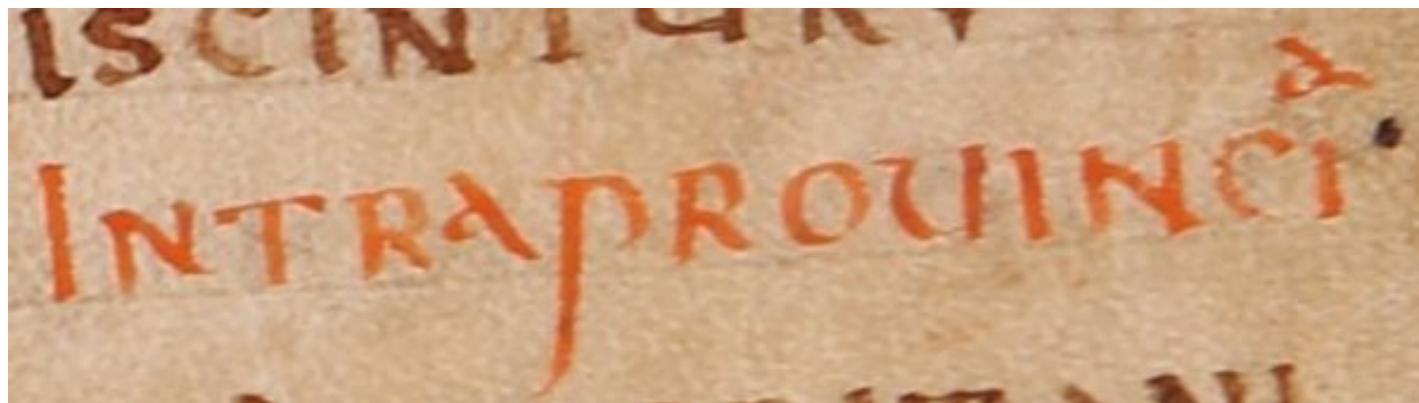
    <!-- fill in the MS sections as <msItem>s -->

    <msItem>
      <locus from="1r" to="9r">1r - 9r</locus>
      <title type="work">Homily I</title>
    </msItem>

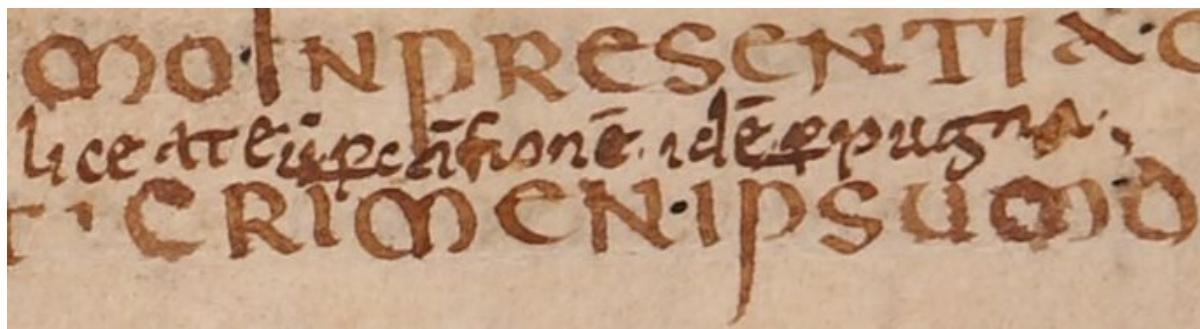
  </msContents>

  <physDesc>
    <objectDesc form="book">
      <supportDesc material="parchment">
        <extent>
          <measure type="leaves">The codex is composed by ff. III (paper, modern); 169 (of which 5 where added in the IX c.); II (paper, XIX c.); III (paper, modern).</measure>
          <measure type="dimensions">Ca. 250 x 180 mm.</measure>
          <measure type="lines">21-23 lines of text per page.</measure>
        </extent>
        <collation>
```

Descrizione del manoscritto



<w><hi rend="cap1.5 empty">S</hi>i</w>
<w>quis</w>
<w>inimicus</w>
<w>intra</w>
<w>pro<c corresp="#v">u</c>inci<add place="above" rend="smaller">a</add></w>



```
<add place="above" hand="#scribe2" >
    <w>liceat</w> <w><choice>
        <abbr><am><g ref="#umacr"/></am></abbr>
        <expan>eu<ex>m</ex></expan>
    </choice>
    </w> <w><choice>
        <abbr><am><g ref="#pbardes"/></am></abbr>
        <expan>p<ex>er</ex></expan>
    </choice></w> <w><choice>
        <abbr>c<am><g ref="#amacr"/></am>ion<am><g ref="#emacr"/></am></abbr>
        <expan><choice>
            <orig>ca<ex>m</ex>fione<ex>m</ex></orig>
            <reg>ca<ex>m</ex>phione<ex>m</ex></reg>
        </choice></expan>
    </choice></w><reg>, </reg> <w>id</w> <w><choice>
        <abbr><g ref="#emacr"/></abbr>
        <expan>e<ex>st</ex></expan>
    </choice></w> <w><choice>
        <abbr><g ref="#pbardes"/></abbr>
        <expan>p<ex>er</ex></expan>
    </choice></w> <w><choice>
        <sic>pugna</sic>
        <corr>pugnam</corr>
    </choice></w><reg>, </reg>
</add>
```

Aggiunte scribali



```
<w>d<subst><del rend="overwrite">o</del><add>u</add></subst>cem</w>
```

Dal Vercelli Book (inglese antico, dialetto tardo sassone occidentale, X secolo ca.):

<del rend="erasure">fffore wæron

<del rend="dot"><g ref="#aeligddot"/>



periculū

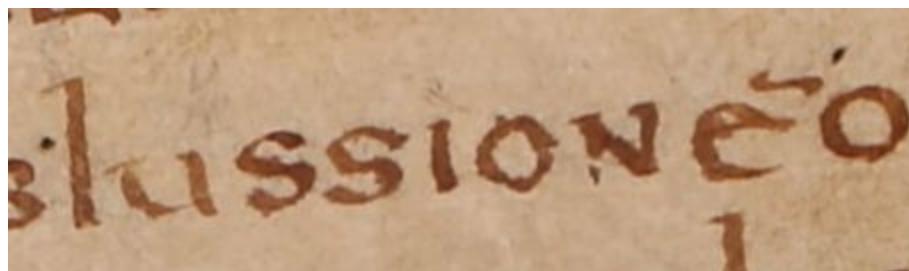
<w>

<choice>

<abbr>pericul<am><g ref="#umacr"/></am></abbr>

<expan>periculu<ex>m</ex></expan>

</choice></w>



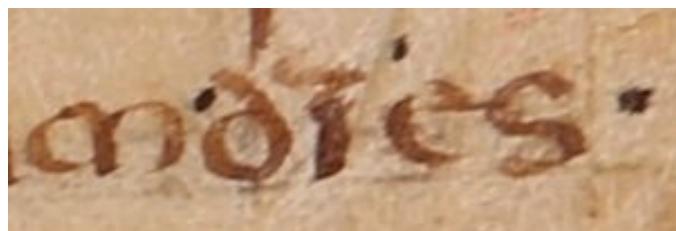
iussionēō

<w><choice>

<abbr>iussion<am><g ref="#emacr"/></am></abbr>

<expan>iussione<ex>m</ex></expan>

</choice></w>



dēi

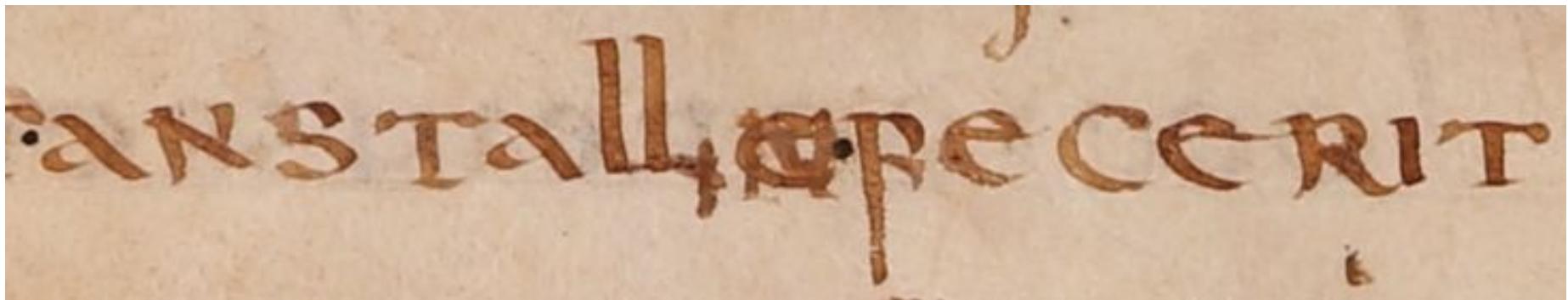
<w>

<choice>

<abbr>d<am><g ref="#imacr"/></am></abbr>

<expan>d<ex>e</ex>i</expan>

</choice></w>



```
<term xml:lang="lng" ref="#astalin" xml:id="CV188_term_astalin_07" corresp="#CV188_gloss_astalin_07">
  <w>
    <choice>
      <orig>anstall<add place="below">i</add><subst>
        <del rend="overwrite">c</del>
        <add>n</add>
      </subst></orig>
      <reg>astalin</reg>
    </choice>
  </w>
</term>
```

Si quis in concilio uel in quol^e
bit c^onventum scandalum comi
serit non ientos solidus sit culpa.

<w>

<choice>

<sic>quole<lb n="11" xml:id="CV188_lb_019v_11_sic"/>bit</sic>

<corr resp="RRDT">quoli<lb n="11" xml:id="CV188_lb_019v_11_corr"/>bet</corr>

</choice></w>

<w>c<space extent="1"/>on<c corresp="#v">u</c>entum</w><l-- conventu -->



<p>
<lb xml:id="CV188_lb_018v_05" n="05"/>
<label n="1"><num><w>I</w></num></label>
<seg rend="cap">
 <w><hi rend="decorated">S</hi>i</w>
 <w>q<c corresp="#u">v</c>is</w>
 <w>homin<c corresp="#u">v</c>m</w>
<lb xml:id="CV188_lb_018v_06" n="06"/><w>contra</w>



```
<p>
<lb xml:id="CV188_lb_018v_10" n="10" type="rubric"/>
<label n="2"><num></w>II</num></label>
<w><hi rend="cap1.2">S</hi>i</w>
<w>quis</w>
<w>cum</w>
<w>rege</w>
<w>de</w>
<w>morte</w>
<w>alte<lb xml:id="CV188_lb_018v_11" n="11"/>rius</w>
<w>consilia<c corresp="#v">u</c>erit</w><reg>, </reg>
<w>aut</w>
<w>hominem</w>
<lb xml:id="CV188_lb_018v_12" n="12"/><w>per</w>
<w>ipsius</w>
<w><choice>
    <abbr>iussion<am><g ref="#emacr"/></am></abbr>
    <expan>iussione<ex>m</ex></expan>
</choice></w>
<w>occiserit</w><reg>, </reg><!-- occiderit -->
<w>in</w>
<lb xml:id="CV188_lb_018v_13" n="13"/><w>nullo</w>
<w>sit</w>
<w>culpa<c corresp="#v">u</c>elis</w><reg>, </reg><!-- culpabilis -->
<w>nec</w>
<w>ille</w>
<w>nec</w>

<pb n="019r" xml:id="CV188_folio_019r" />
```

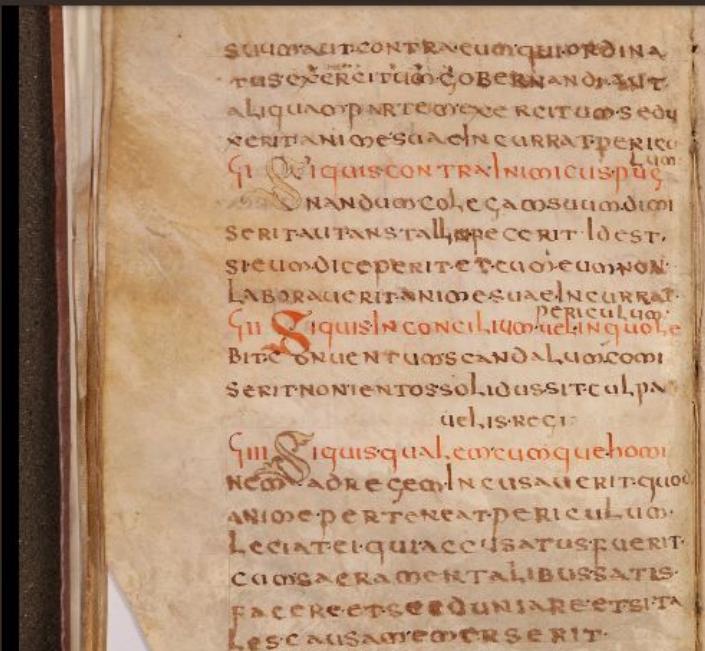
Marcatura delle singole parole



```
<term xml:lang="lng" ref="#astalin" xml:id="CV188_term_astalin_07" corresp="#CV188_gloss_astalin_07">
  <w>
    <choice>
      <orig>anstall<add place="below">i</add><subst>
        <del rend="overwrite">c</del>
        <add>n</add>
      </subst></orig>
      <reg>astalin</reg>
    </choice>
  </w>
</term>
<w>fecerit</w><reg> ,</reg>
<gloss target="#CV188_term_astalin_07" xml:id="CV188_gloss_astalin_07">
  <w>id</w>
  <w>est</w>
  <lb n="08" xml:id="CV188_lb_019v_08" /><w>si</w>
  <w>eum</w>
  <w>diceperit</w>
  <w><damage>et</damage></w>
  <w>cum</w>
  <w>eum</w>
  <w>non</w>
  <lb n="09" xml:id="CV188_lb_019v_09" /><w>labora<c corresp="#v">u</c>erit</w>
</gloss>
```

Thumbs Magnifier MS Desc

Editctus Rothari 19v Diplomatic



36%

Search Lists No selection A- A+

Thumbs

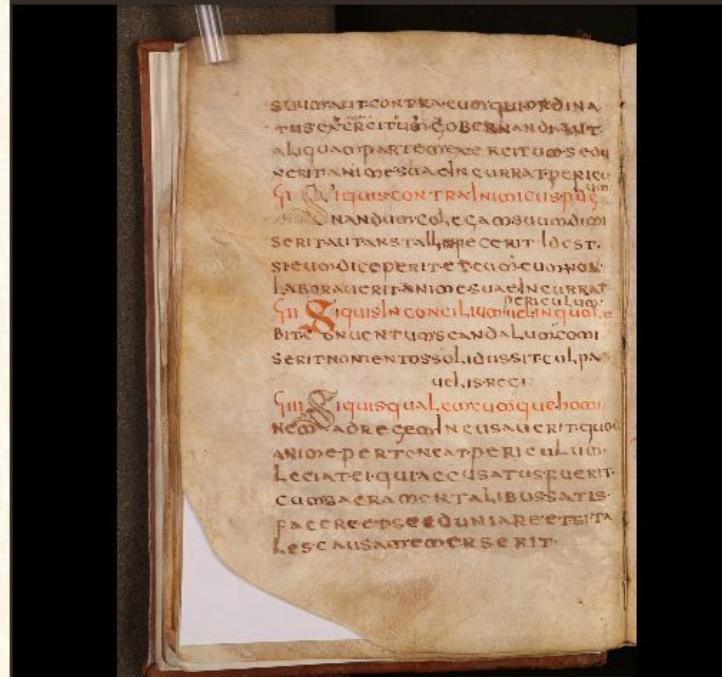
Magnifier

MS Desc

Editctus Rothari

19v

Interpretative



27%

- 01 suum aut contra eum, qui ordina
02 tus exercitum gubernandi, aut
03 aliquam partem exercitum sedu
04 xerit, animae suea incurrat periculum.
05 VII Si quis contra inimicus pug
06 nandum colegam suum dimi
07 serit aut astalin fecerit, id est
08 si eum diceperit et cum eum non
09 laborauerit, animae suea incurrat periculum.
10 VIII Si quis in concilium vel in auole
11 bit conuentum scandal
12 serit, non ientos solidus
13 uelis regi.
14 VIII Si quis qualemcu

- Term
- Gloss
- Select All
- Clear selection

Search

Lists

Multi selection

A-

A+

Thumbs

Magnifier

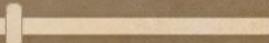
MS Desc

Editus Rothari

20v

Interpretative

MORTUOS NON FUERIT CONPONAT PUNI
QUISQUE SOLEDOS SUI GENTES SI EXIP
SI CONSILIUAM MORTUOS FUERIT CON
PONAT IPSUM MORTUOS SI CUTA
PRAE CINTUOS FUERIT IDES TUA RICHILO
XII. **I**DO **U**NO AUT **T**RES HOMINES SISSUS
AM PLUSTRIBI HUMICIDIA OMNIA PENET
TRAUERIT ET VOLUCRENT SE AD UNA
RE CTHUNUM CONPONANT SI CUT
AD PRECINTUOS FUERIT SITE IS LICEN
TIAM AD UNANIDIETIS ALIQUIIS DEIP
SISS SUBTRAXERIT NON POTUERIT
SEPARIFICARE SI CUT LECHABIT
QUOD NEE PLAEXA OMNIS FERIT ALIIP
SUM HOMINEM QUI OCCISUS EST NON
FECISSIT TUNCS IT CULPA UELIS
SICUT ET ALIQUI CONPOSUIT ET SI
SEPARIFICAUERIT SIT ABSOLUTUS
APENAHUORI EIDU SIT NOCENDU CON
SILIO FUERIT CONPONAT UT SUPRA



36%

List of words

List of glosses

A B C D E F G H I J K L M N O P Q R S T U V W X Y

aldium

Fol. 31r - Doc. Edictus Rothari

Fol. 31v - Doc. Edictus Rothari

Fol. 32r - Doc. Edictus Rothari

Fol. 32v - Doc. Edictus Rothari

Fol. 33v - Doc. Edictus Rothari

aldius

angargathungi

Search

Lists

Multi selection

A-

A+



The digital facsimile



The spread of the digital facsimile

- a natural evolution of the photographic facsimile
- usually only images with accompanying metadata and general information about the manuscript
- the path followed by many famous libraries
 - British Library <https://www.bl.uk/manuscripts/>
 - Vatican Apostolic Library <https://www.vaticanlibrary.va/>
 - Oxford Bodleian Library <https://digital.bodleian.ox.ac.uk/>
- and by others that have put many MSS online
 - e-codices <https://www.e-codices.unifr.ch/en>

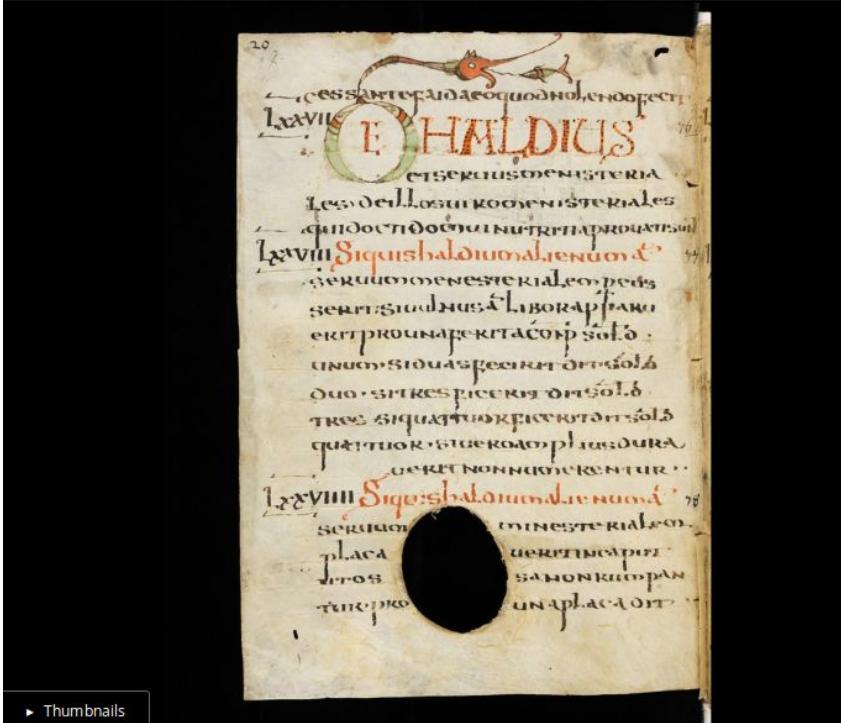


Pagina: 20



Visualizzazione: Aprire in Mirador

Sequenza: versione restaurata (immagini d'aprile 2009)



▶ Thumbnails

Informazioni di base

Come citare

Diritti

Descrizione



St. Gallen, Stiftsbibliothek, Cod. Sang. 730

Pergamena · 88 pp. · 20.5 x 14 cm · Bobbio (?) · 670/680

Edictum Rothari (Veterum Fragmentorum Tomus III)

Verso la visione d'insieme

Visualizza sequenza: versione restaurata (immagini d'aprile 2009)

Descrizione della sequenza: L'*Edictum Rothari* è stato restaurato nel 2008 dal restauratore Martin Streb. I materiali accessori inseriti per ricostruire lo stato originale del manoscritto sono stati tolti, tra questi anche delle fotografie di frammenti di Zurigo (Archivio di Stato e Biblioteca centrale) e di Karlsruhe (Badische Landesbibliothek). E' prevista una ricostruzione digitale con tutte le parti rimaste, per la collezione "sine loco".

Lingua: Latino

Descrizione breve: Editto di Rotari, conservato in frammenti. È la più antica copia della raccolta di leggi dei Longobardi redatta e promulgata da re Rotari (636-652) nel 643. Realizzata probabilmente a Bobbio e databile agli anni 670/680, è conservata in frammenti divisi tra la biblioteca abbaziale di San Gallo, la Badische Landesbibliothek



Text-image relationship



Text-image relationship

Text-image relationship

- **targeted link** (hot-spot): a specific area of the image is highlighted in such a way that, by interacting with it, information is displayed such as editorial notes, improved version of a detail, comment on the text, etc.
- **generalized link**: all the text of the edition is placed in direct relation with the corresponding images, or parts of the image, so as to be able to easily access the corresponding portion of the image starting from the text, and vice versa

Generalized link

- page level link: an aid in navigation, especially in the case of large manuscripts (or editions comprising more than one manuscript)
- line / word level link: an important teaching aid, the student can follow the text line by line or word by word
 - support for other features, especially word search or palaeographical details, based on encoding
 - the more granular the linking type, the greater the coding burden

Technical implementation: HTML + JavaScript

- image maps method use here: <http://web.uvic.ca/hrd/lydgate/>
- ingenious and reasonably efficient, but has several disadvantages:
 - quite complicated to make
 - the dimensions of the "stripes" are fixed: strict requirement, they may not correspond exactly to the stripes
 - hot-spot support needs to be implemented separately
 - everything is done at the presentation level: the XML encoding does not carry the necessary information

And so toke a swerd ful sharpe wher o gromond
 And y with all the rof hir herd in thome
 So haugh y vngydane wrope w agynd welcomd
 On them i falleid don y gylde penis
 To stamidre folke for hit os yor ordignis
 With yor de fute oy folke to agyntz atbute
 God at y lade yor malice can agyntz.
That som boyle of phedra do record
 That she assaynd and confusid of y dede
 hymur hir folke xp ful highe with accord
 So haugh ful stamidre can folke quyte s mado
 Wher for y comysl em man take hele
 In fute matore as stond in no certen
 From hasty lauerage yor tongis to restremo
De casu Cizars dutis Regis labyn
Hmonyng storis woful for to rede
 Al be sprent with terys in his face
 Ful sodenlie Iohn Bochas can tak hede
 Amyddys of p^e prete 3i3ara cam to place
 And haugh y fortun can eke to manesse
 This proude duke ful myghty & notable
 Of kyng labyn callid p^e constable
Of his ofte ledere and governere
 To Israel verrey mortall foo
 With peple he roode lyke a conyndrome
 And wher y^t er he meny did go
 Tho eth quode yo peple drie hym b
 Pledo from his face wher he cam aferro
 w hundred maynes he haddes for y warre
Cstrongly enarmyd for yo warre

Partial Transcription of John Lydgate's "Fall of Princes"

[Header information suppressed] [Back to top](#)

12/1/1
 12/1/2
 12/1/3
 12/1/4
 12/1/5
 12/1/6
 12/1/7
 12/1/8
 12/1/9
 12/1/10
 12/1/11
 12/1/12
 12/1/13
 12/1/14
 12/1/15
 12/1/16
 12/1/17
 12/1/18
 12/1/19
 12/1/20
 12/1/21
 12/1/22
 12/1/23
 12/1/24
 12/1/25
 12/1/26

De casu Cizars dutis Regis labin

→ **A**mong op^e stories woful for to rede
 Al be sprent with terys in his face
 Ful sodenlie Iohn Bochas can tak hede
 Amyddys of p^e prete 3i3ara cam to place
 And haugh y fortun can eke to manesse
 This proude duke ful myghty & notable
 Of kyng labyn callid p^e constable
 Of his ofte ledere and governere

The TEI P5 digital facsimile

- the TEI encoding schemas version P5 (2007) introduce numerous improvements regarding the management and transcription of manuscripts
- among these the new *Digital Facsimiles* section in chapter 11 *Representation of Primary Sources*:
 - <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/PH.html>
 - previous work in the same area:
<http://users.ox.ac.uk/~lou/wip/digfax.html>

New `facs` attribute

- including the `transcr` module in the TEI encoding scheme makes a new global attribute available:
- **facs** (`facsimile`) points to all or part of an image which corresponds with the content of the element
- this attribute can be used in any element to associate the content of the same to an image, directly or by linking to `<graphic>` or `<surface>` or other elements:

```
<p n="1" facs="#para1_img"> <head facs="#head_img">  
<pb facs="page1.jpg"/>
```

New elements and `start` attribute

- `<facsimile>` contains a representation of some written source in the form of a set of images rather than as transcribed or encoded text
- `<surface>` defines a written surface in terms of a rectangular coordinate space, optionally grouping one or more graphic representations of that space, and rectangular zones of interest within it
- `<zone>` defines a rectangular area contained within a surface element
- `@start` points to an element holding the starting position of the text corresponding to the inscribed part of the surface

The <facsimile> element

- the <facsimile> element is of a structural type and is placed at the same level as <text> or even as an alternative to the latter
- when the *transcr* module is added to the encoding scheme it is possible to choose between:
 - a <teiHeader> and a <facsimile>
 - a <teiHeader> and a <text>
 - a <teiHeader>, a <facsimile> and a <text>
- this allows great flexibility:
 - simplest case: digital facsimile with images of the ms
 - digital facsimile with text transcription
 - digital facsimile with transcription and text-image link

Digital facsimile with images only



```
<TEI>
  <teiHeader>
    <!--...-->
  </teiHeader>
  <facsimile>
    <graphic url="page1.png"/>
    <graphic url="page2.png"/>
    <graphic url="page3.png"/>
    <graphic url="page4.png"/>
  </facsimile>
</TEI>
```



```
<TEI>
  <teiHeader>
    <!--...-->
  </teiHeader>
  <text>
    <pb facs="page1.png"/>
      <!-- inserire qui il testo di pagina 1 -->
    <pb facs="page2.png"/>
      <!-- inserire qui il testo di pagina 2 -->
  </text>
</TEI>
```

Disadvantages of <text> with no <facsimile>

- on the basis of this encoding it is possible to generate an HMTL page divided into two panes (image and text) for example using an XSLT stylesheet
- disadvantages of this solution:
 - no non-page level text-image links
 - it is not possible to locate particular areas of the images
 - pointers to images are scattered throughout the document (because there is no <facsimile> element)

Using both <text> and <facsimile>

- the most effective and flexible solution is the a parallel transcription based on <facsimile> and <text>
- using of <surface> and <zone> inside <facsimile>
 - to define image areas
 - to link the transcript text to those areas and/or subimages
- the image areas are identified by means of a Cartesian coordinate system and recorded as values of the following attributes of <surface> and <zone>:
 - **ulx, uly** x and y coordinates of the upper left corner
 - **lrx, lry** x and y coordinates of the lower right corner

Using both <text> and <facsimile>

- the reference system is common for all elements that record coordinates
- the only limitation is given by the fact that the areas defined with the attributes indicated above can only be of the rectangular type
- in the case of polygonal areas, it is still possible to carry out a transcription using the TEI schemes
 - new **points** attribute for polygonal areas
- the viewing software must highlight these areas on the screen and connect with the text

Using <surface>

- <surface> identifies the written surface of an image

```
<surface ulx="54" uly="20" lrx="400" lry="280">  
    <graphic url="page1.png"/> </surface>
```

- can contain more than one <graphic> element

```
<surface ulx="54" uly="20" lrx="400" lry="280">  
    <graphic url="page1-highRes.png"/>  
    <graphic url="page1-lowRes.png"/>  
</surface>
```

- in addition to / instead of <graphic> it can contain one or more <zone>s
- as many <surface>s inside the <facsimile> as there are folios

Using <zone>

- <zone> defines a specific area of the image using the same coordinate system as <surface>, inside or outside or overlapping the <surface>:

```
<surface ulx="54" uly="20" lrx="500" lry="321">
    <zone ulx="54" uly="20" lrx="400" lry="280">
        <graphic url="paragraph.png"/></zone>
        <note>first page</note>
    </surface>
```

- or a smaller portion (useful for creating a hot-spot):

```
<zone ulx="120" uly="48" lrx="143" lry="56">
    <graphic url="gloss.png"/>
    <note>Scribe gloss</note>
</zone>
```

The Lindisfarne Gospels



- the first step is the identification of areas of interest
- main text, glosses, decorated initials, etc.
- each of these must be converted into a <zone>

<zone> elements within a <surface>



```
<facsimile xml:id="imtAnnotatedImage">
  <surface>
    <graphic height="1797px" url="LindisfarneFol27rIncipitMatt.jpg"
width="1266px"/>
    <zone lrx="1268" lry="1797" rend="visible" rendition="surface"
ulx="0" uly="-4" xml:id="imtArea_0"/>
    <zone lrx="1267" lry="450" rend="visible" rendition="zone"
ulx="1202" uly="356" xml:id="imtArea_1"/>
    <zone lrx="1050" lry="792" rend="visible" rendition="zone"
ulx="81" uly="30" xml:id="imtArea_3"/>
    <zone lrx="1190" lry="154" rend="visible" rendition="zone"
ulx="503" uly="48" xml:id="imtArea_4"/>
    <zone lrx="1184" lry="412" rend="visible" rendition="zone"
ulx="1116" uly="353" xml:id="imtArea_6"/>
    <zone lrx="86" lry="606" rend="visible" rendition="zone" ulx="2"
uly="478" xml:id="imtArea_5"/>
    <zone lrx="995" lry="366" rend="visible" rendition="zone"
ulx="694" uly="321" xml:id="imtArea_6"/>
  </surface>
</facsimile>
```

Linking text to <surface> and <zone> elements

- to link the **text** of the transcript to the corresponding areas of the image:
 - assign a unique identifier to each element of the facsimile using the `xml:id` attribute
 - use the `facs` attribute in text elements to specify the id of the corresponding <surface> and <zone> elements
- to link **image areas** to corresponding text elements:
 - assign a unique identifier to each element of the transcript using the `xml:id` attribute
 - use the `start` attribute (also `corresp`) in the <surface> and <zone> elements to specify the id of the text elements

Bi-directional example



```
<facsimile>
  <surface xml:id="page1" start="#page_1" ulx="0" uly="0" lrx="500"
lry="321">
    <graphic url="page1.png"/>
    <zone xml:id="line1" ulx="50" uly="80" lrx="200" lry="321">
        <graphic url="line1.png"/>
        <note>First page.</note>
    </zone>
    <zone xml:id="det1" ulx="120" uly="48" lrx="143" lry="56">
        <graphic url="gloss.png"/>
        <note>Scribe gloss.</note>
    </zone>
  </surface>
</facsimile>
<text>
  <body>
    <div>
      <pb facs="#page1" n="1" xml:id="page 1"/>
      <p>Lorem ipsum ... <gloss facs="#det1">semper</gloss></p>
    </div>
  </body>
</text>
```

How to calculate coordinates

- once decided the linking type, it is essential to use the correct tool to define image areas by means of cartesian coordinates, and to include them as values for the relevant TEI elements
- fortunately there are several programs to help us with this task:
 - TEI Zoner: <http://teicat.huma-num.fr/zoner/index.php>
 - Image Markup Tool:
https://hcmc.uvic.ca/~mholmes/image_markup/
 - Oxygen XML Editor Image Map:
<https://www.oxygenxml.com/doc/versions/25.0/ug-editor/topics/image-map-tei.html>

TEI Zoner

- developed by Chris Sparks, University of London
- Web application that allows the user to define rectangular or polygonal areas and annotate them
- requires no knowledge of XML or TEI, the output can be integrated into a TEI document
- rather user friendly user interface, [recently improved](#) to manage individual areas (e.g., to copy or delete them in case of error)
- produced data are compatible with TEI P5 schemas, recently made compatible with IMT (categories and annotations), but it is not possible to save a project

Annotations Categories Fields

Show TEI Draw Rectangle Clear points Clear shapes Delete

1 x:229px (19%)y:294px (16%)x,y:229, 294 ?

Text Input

Current annotation: _blank

Category: detail

Fields:

title

_blank

_blank

_blank

+ Add annotation

dam dæg dyrhæf manna ge hylcum sce lif. spa in
hæfna pieſ hyhðo fya on helle ritū fya he nu hæf onpopulde
ge aqingan pile tilgh pſtu forhan þame us god forgrif
hæt pe his þa halgan lare gehyrni spa pe mudydon hæt pe
yanþe bethan þe rehun fori ban lifum in salha populda
populd to ridan fforpe mid fædri ƿimð runa ƿimð ham halgan
gafe Incestre aman; lægi spel to fþylcne tide

MEN da lafstan hrynt halige dægar ƿastlice ƿifru
raplum lacadomlice ƿe miele nyð heaſfe habbadý
ƿe dat gærnlice gehlycni ƿifru hylle he pe þane tide
benigh he us misyt god for his mild heortnare ƿaymed hafas
he pe da halgan godis aƿe ƿa godspellican lare inupum mode

TEI Zoner: area selection on image

Annotations Categories Fields

Hide TEI

Draw Rectangle

Clear points Clear shapes Delete



1

x:255px (21%)y:304px (17%)x,y:255, 304

?

Text Input

Current annotation: blank

Category: detail

Fields:

title

_blank



_blank



_blank



+ Add annotation

TEI Markup

JSON XML zones XML annotations/text XML categories IMT XML

```
<graphic url='Vercelli-Book_076V_S_300dpi.jpg' width='1200' height='1800' />  
  
<zone xml:id='imtArea_0' points=' 637,666 374,666 374,620 637,620 ' rendition="line" />  
  
<zone xml:id='imtArea_1' points=' 245,637 1057,637 1057,730 245,730 ' rendition="line" />  
  
<zone xml:id='imtArea_2' points=' 249,769 384,769 384,645 249,645 ' rendition="detail" />
```

I am summae scripturam Bonum paupile pe pe pape tide
benevolentie us misericordia populi his mild heartenare grymed hafas
per pe da halgan godis ape Iba godspellican lige in ipum mode

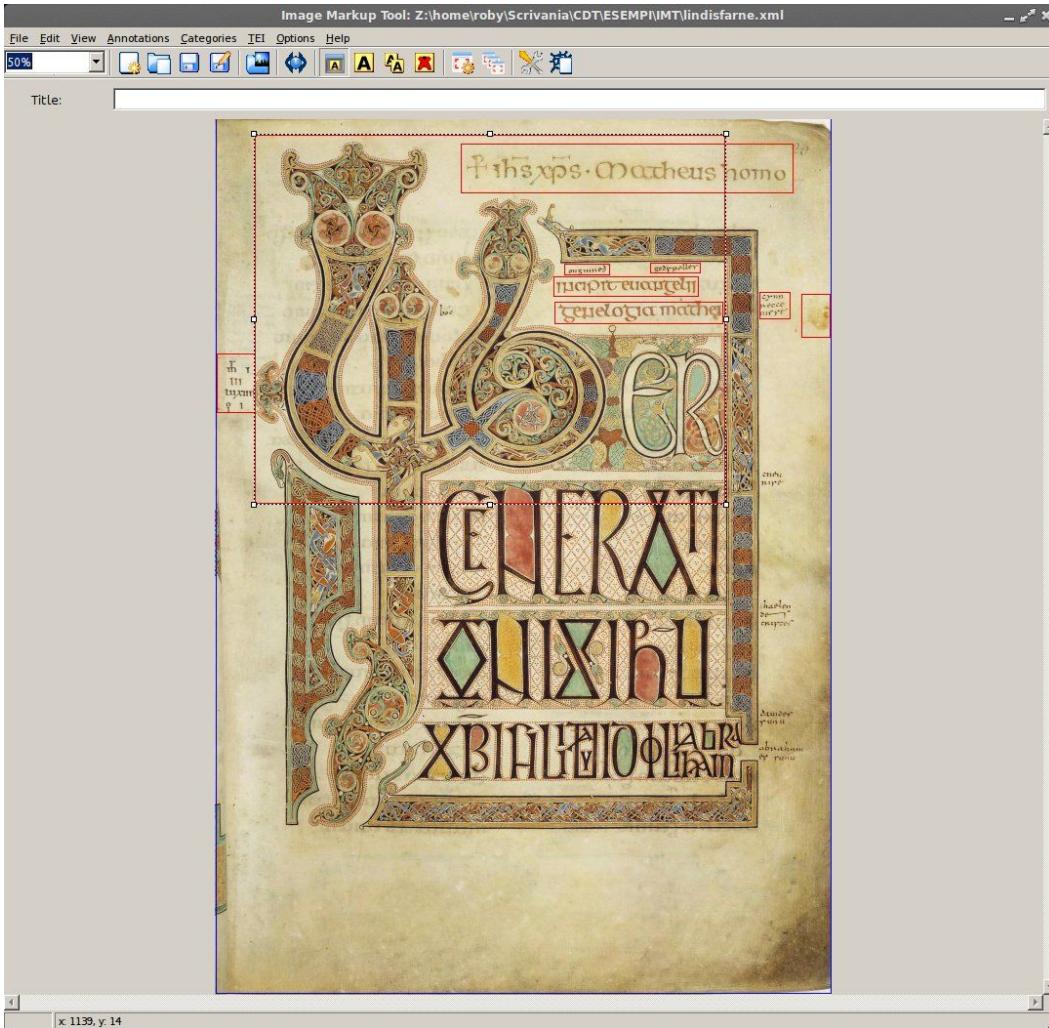
TEI Zoner: the corresponding TEI markup

The UVic Image Markup Tool

- developed by Martin Holmes, University of Victoria (CA)
- an open source tool that allows the user to define rectangular areas and annotate them, also according to categories
- requires no knowledge of XML or TEI, the output can be integrated into a TEI document
- saved data are saved according to a TEI P5 schema
- very user friendly user interface, category and id management
- a Windows program also usable under Linux (with Wine) and MacOS X (with DarWine)
- unfortunately development has been discontinued

The UVic Image Markup Tool

- all annotations are saved in the project file, an XML document conforming to the TEI P5 standard
- a web view of the document can be generated at any time
- the software will automatically produce HTML and Javascript code
 - the text-image link is bidirectional
 - large images will be reduced in size
 - a zoom function will be added for annotated areas
- the web view can be saved and customized
- sample project: <http://mariage.uvic.ca/>



Annotations by category

A **A** **A** **D** **D** **C** **C** **K**

All
zone
surface

Category properties...

Annotations

Show all annotations in this category

- Area marcata come surface
- Area danneggiata marcata come zone
- Decorazione zoomorfa
- Intestazione
- [Annotation title]
- [Annotation title]
- [Annotation title]
- [Annotation title]

Annotation details

Category:

Annotation title: ID...

Annotation text:

OK Cancel



Annotations	
zone	Area danneggiata marcata come zone
	Decorazione zoomorfa
Intestazione	[Annotation title] [Annotation title] [Annotation title] [Annotation title] [Annotation title] [Annotation title]
surface	Area marcata come surface

Using the Image Markup Tool

- very effective tool, but
 - limited to single images for now
 - no support for editing XML code
 - the XML files obtained must be integrated by hand with the transcript files
- IMT then needs to be integrated into an environment that includes other working tools:
 - a good XML editor
 - a software for visualizing the edition
- the alternative is to use an integrated product (Oxygen, TextGrid)

Oxygen Image Maps in TEI

- all the latest versions of Oxygen offer support for <zone> definition within a <facsimile>
- advantages:
 - seamless integration into the TEI document
 - support for polygonal areas
- disadvantages:
 - proprietary software
 - not so friendly to use



100%

1

De Geometricie.

49

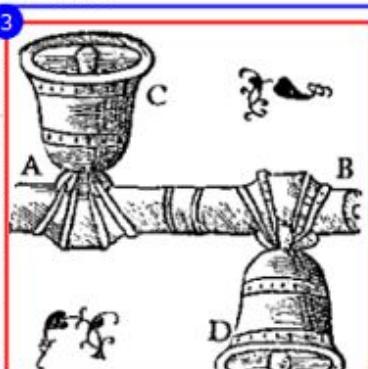
¶ DV SON ET ACCORD DES CLOCHEZ ET
des alleures des cheualx, chariotz & charges, des fontaines: &
encyclie du monde, & de la dimension du corps humain.

Chapitre septiesme.

¶ Le son & accord des cloches pendans en vng mesme
axe, est faict en contraires parties.

2

Les cloches ont quasi figures de rondes pyramides imperfaictes &
irregulieres: & leur accord se fait par regle geometrique. Comme
les deux cloches C & D sont pendans i vng mesme axe
ou effieu A B: ie dis que leur accord se fera en contraires parties,



C

A

B

D

There are shapes entirely covered by bigger shapes.

Type: RECT [80, 63, 419, 70]

ID: B49rHead

?

#	Type	ID:
1	Rectangle	B49rHead
2	Rectangle	B49rPara2
3	Rectangle	B49rFig1
4	Rectangle	B49rW457

OK

Cancel



The embedded transcription



The embedded transcription

- a method somewhere between a digital facsimile and an image-based edition is that of embedded transcription:

<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/PH.html#PHZLAB>

- important differences compared to <text>based parallel transcription:
 - text is considered a supporting aid, focus on images (e.g. physical arrangement of parts)
 - in fact the encoding model of <line> is limited
 - fewer transcription options to avoid hierarchy conflicts

Facsimile based on embedded transcription

- use of the `<sourceDoc>` element on the same hierarchical level of, and possibly as an alternative to, `<facsimile>` and `<text>`
- use of `<surface>` and `<zone>` in a manner similar to what was seen above
- `<zone>` contains a series of `<line>` elements corresponding to lines of text:

```
<zone ulx="20" uly="40" lrx="120" lry="180">
    <line>first line of transcription</line>
    <line>second line of transcription</line>
</zone>
```

```
<sourceDoc>
  <surfaceGrp n="South_Wall">
    <surface xml:id="EETS.29">
      <zone>
        <graphic url="img_0479.JPG"/>
        <line><orig>Now <damage>in the</damage> name of ou<damage>r</damage>e
          l<damage>or</damage>d ih<ex>esu</ex>s</orig></line>
        <line><orig>of right hole herte <ex>&amp;</ex> in our <damage>b</damage>est
          entent</orig></line>
        <line><orig>ou<damage>r</damage>
          <damage>l</damage>yf reme<ex>m</ex>bryng froward and
          vicious</orig></line>
        <line><orig>a<damage>y</damage> co<damage>n</damage>t<damage>ra</damage>ye to
          the commandement</orig></line>
        <line><orig>of crist i<ex>hes</ex><damage>u</damage> now wy<damage>th</damage>
          aviseme<damage>nt</damage></orig></line>
        <line><orig>the l<damage>ord beseching of mercy and pete</damage></orig></line>
        <line><orig>our you<damage>th</damage>e <ex>&amp;</ex> a<damage>g</damage>e
          tha<damage>t</damage> we have <damage>myspent</damage></orig></line>
        <line><orig>wyt<damage>h this woord mercy knelyng o</damage>n
          o<damage>ur</damage> kne</orig></line>
      </zone>
    </surface>
    <surface />
    <surface />
  </surfaceGrp>
<sourceDoc>
```

Test for rendering sourceDoc

Now in the name of oure lord ihesus
of right hole herte & in our best entent
our lyf remembryng froward and vicious
ay contrarye to the commandement
of crist ihesu now wyth avisement
the lord beseching of mercy and pete
our youthe & age that we have myspent
wyth this woord mercy knelyng on our kne

caesar adsum iam forte

[TEI Home](#)

Matthew Davis. Date: 2013-11-13





Character encoding and the TEI gaiji module



Character encoding

- text transcription by visible graphemes on the medium that preserves it is the basic level of the encoding process
- the minimum unit is the character, we distinguish between
 - **character** = an abstract entity that has a number of fundamental traits
 - **glyph** = the graphic realization of a given character

A A A A A A A A A A A A A A A A

- a similar distinction to that between phoneme and phonetic sound, grapheme and glyph
- text encoding occurs through character encoding, but this process is trivial only in appearance

Computers, numbers and characters

- the computer can only interpret and process numbers
- "characters" are just other numerical data
- "character encoding" means assigning a number to a character so that it can be identified
- for example, the character A (ASCII table) corresponds to the
 - 65 decimal system
 - 41 hexadecimal system
 - 0100001 binary system
- character set (font) = a collection of glyphs associated with the corresponding characters by means of numeric codes

The ASCII standard

- early encoding systems were quantitatively limited (ASCII, EBCDIC)
- the American Standard Code for Information Interchange (ASCII) standard is still used (for retrocompatibility mainly)
 - Original ASCII 7 bits 128 characters
 - Extended ASCII 8 bits 256 characters
- extended ASCII barely covers the Latin alphabet and a little more
- this limit is exceeded with ISO-8859-* extensions (Latin, Greek, Cyrillic, etc.) but at the cost of interchange
 - the same code is assigned to different characters → specific fonts
 - usable only within the same writing system or you have to change font

Hex	Dec	Char	Hex	Dec	Char	Hex	Dec	Char	Hex	Dec	Char
0x00	0	NULL null	0x20	32	Space	0x40	64	@	0x60	96	`
0x01	1	SOH Start of heading	0x21	33	!	0x41	65	A	0x61	97	a
0x02	2	STX Start of text	0x22	34	"	0x42	66	B	0x62	98	b
0x03	3	ETX End of text	0x23	35	#	0x43	67	C	0x63	99	c
0x04	4	EOT End of transmission	0x24	36	\$	0x44	68	D	0x64	100	d
0x05	5	ENQ Enquiry	0x25	37	%	0x45	69	E	0x65	101	e
0x06	6	ACK Acknowledge	0x26	38	&	0x46	70	F	0x66	102	f
0x07	7	BELL Bell	0x27	39	'	0x47	71	G	0x67	103	g
0x08	8	BS Backspace	0x28	40	(0x48	72	H	0x68	104	h
0x09	9	TAB Horizontal tab	0x29	41)	0x49	73	I	0x69	105	i
0x0A	10	LF New line	0x2A	42	*	0x4A	74	J	0x6A	106	j
0x0B	11	VT Vertical tab	0x2B	43	+	0x4B	75	K	0x6B	107	k
0x0C	12	FF Form Feed	0x2C	44	,	0x4C	76	L	0x6C	108	l
0x0D	13	CR Carriage return	0x2D	45	-	0x4D	77	M	0x6D	109	m
0x0E	14	SO Shift out	0x2E	46	.	0x4E	78	N	0x6E	110	n
0x0F	15	SI Shift in	0x2F	47	/	0x4F	79	O	0x6F	111	o
0x10	16	DLE Data link escape	0x30	48	0	0x50	80	P	0x70	112	p
0x11	17	DC1 Device control 1	0x31	49	1	0x51	81	Q	0x71	113	q
0x12	18	DC2 Device control 2	0x32	50	2	0x52	82	R	0x72	114	r
0x13	19	DC3 Device control 3	0x33	51	3	0x53	83	S	0x73	115	s
0x14	20	DC4 Device control 4	0x34	52	4	0x54	84	T	0x74	116	t
0x15	21	NAK Negative ack	0x35	53	5	0x55	85	U	0x75	117	u
0x16	22	SYN Synchronous idle	0x36	54	6	0x56	86	V	0x76	118	v
0x17	23	ETB End transmission block	0x37	55	7	0x57	87	W	0x77	119	w
0x18	24	CAN Cancel	0x38	56	8	0x58	88	X	0x78	120	x
0x19	25	EM End of medium	0x39	57	9	0x59	89	Y	0x79	121	y
0x1A	26	SUB Substitute	0x3A	58	:	0x5A	90	Z	0x7A	122	z
0x1B	27	FSC Escape	0x3B	59	;	0x5B	91	[0x7B	123	{
0x1C	28	FS File separator	0x3C	60	<	0x5C	92	\	0x7C	124	
0x1D	29	GS Group separator	0x3D	61	=	0x5D	93]	0x7D	125	}
0x1E	30	RS Record separator	0x3E	62	>	0x5E	94	^	0x7E	126	~
0x1F	31	US Unit separator	0x3F	63	?	0x5F	95	_	0x7F	127	DEL

ASCII table (7 bits per character)

Note the amount of control codes that occupy the entire first column:
 ASCII originated as a standard for Bell teleprinters.

Only a few encodings available for the actual characters.

Limitations of the ISO-8859 extensions

- besides the inevitable interchange problems, ISO-8859 encodings are still limited, particularly for ancient or medieval writing systems
- a very large number of characters are not present, therefore encoding texts becomes very difficult
- solution #1: creation of *ad hoc* fonts (TTF fonts) so that all necessary characters are available
- disadvantages:
 - not always cheap or easy to implement
 - serious interchange problems (based on no standard)
 - impractical for text dissemination via WWW (you would need *that* font)

Limitations of the ISO-8859 extensions

- solution #2 from the *Dictionary of Old English* corpus:

[005900 (162)] Ffne swa he cwæde: Micel is gefea,
he ne bið Iudeum anum seald, ac he bið eallum þeo
willað.

- the corresponding HTML code:

```
fne swa he cwæde: Micel is gefea, [...]
```

- the problem is not exclusively aesthetic: it is actually impossible to perform text searches if no characters are employed!
- an advanced encoding is needed for non-standard or special characters (variants), one valid for all writing systems

The Unicode standard

- these problems drove toward the definition of a truly universal standard
- developers from Xerox, Apple and others begin designing it in 1987
- the project gathered support, and in 1991, after the founding of the Unicode consortium, version 1.0 of the standard was published
- the latest version (v. 15.0) dates from September 2022 and includes more than 149,000 characters from various writing systems (Latin, Arabic, Chinese, Tamil, Greek, Cyrillic, etc.)
- website: <https://home.unicode.org/> <https://www.unicode.org/versions/latest/>

Unicode goals

- a standard that is truly "universal, efficient, unambiguous"
- fundamental distinction between character and glyph
- periodically synchronized with the ISO/IEC 10646 standard
- support for recent and older writing systems

"The Unicode Standard is a character coding system designed to support the worldwide interchange, processing, and display of the written texts of the diverse languages and technical disciplines of the modern world. In addition, it supports classical and historical texts of many written languages." (*About the Unicode Standard*, <http://www.unicode.org/standard/standard.html>)

Unicode features

- Unicode is **compatible** with the ASCII standard ("embedded" in the new encoding, ASCII numeric values remain identical)
- Unicode is **extensible**: defined more than 149k characters, available space exceeds one million
- Unicode provides for a **PUA - Private User Area**:
 - unused code points which can be assigned to user-defined characters
 - subsequently implementable in a font
 - they may become part of the standard
- Unicode accepts **requests** from individual scholars or institutions to add new characters

MUFI: The Medieval Unicode Font Initiative

“The Medieval Unicode Font Initiative is a non-profit workgroup of scholars and font designers who would like to see a common solution to a problem felt by many medieval scholars: the encoding and display of special characters in Medieval texts written in the Latin alphabet.”

- new characters entered into PUA and proposed to Unicode
- if accepted the PUA code point is removed in favor of the final one (so watch out! see below)
- 152 accepted fonts in Unicode 5.1
- fonts that implement MUFI are listed on the home page
- website: <https://mufi.info/>

Figure 1-1. Wide ASCII



ASCII/8859-1 Text

A	0100 0001
S	0101 0011
C	0100 0011
I	0100 1001
I	0100 1001
/	0010 1111
8	0011 1000
8	0011 1000
5	0011 0101
9	0011 1001
-	0010 1101
1	0011 0001
	0010 0000
t	0111 0100
e	0110 0101
x	0111 1000
t	0111 0100

Unicode Text

A	0000 0000 0100 0001
S	0000 0000 0101 0011
C	0000 0000 0100 0011
I	0000 0000 0100 1001
I	0000 0000 0100 1001
	0000 0000 0010 0000
天	0101 1001 0010 1001
地	0101 0111 0011 0000
	0000 0000 0010 0000
沮	0000 0110 0011 0011
沮	0000 0110 0100 0100
沮	0000 0110 0010 0111
	0000 0110 0100 0101
	0000 0000 0010 0000
a	0000 0011 1011 0001
\$	0010 0010 0111 0000
Y	0000 0011 1011 0011

Tabella caratteri

File Visualizza Cerca Vai Aiuto

Times New Roman ▾ Grassetto Corsivo 26 - +

Scrittura ▾ Tabella caratteri Dettagli carattere

Iscrizioni parthian	A	B	C	D	E	F	G	H	I	J
Javanese	K	L	M	N	O	P	Q	R	S	T
Kaithi	U	V	W	X	Y	Z	a	b	c	d
Kannada	e	f	g	h	i	j	k	l	m	n
Katakana	o	p	q	r	s	t	u	v	w	x
Kayah Li	y	z	a	o	À	Á	Â	Ã	Ä	Å
Kharoshthi	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
Khitan Small Script	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	Ø	Ù	Ú
Khmer										
Khojki										
Khudawadi										
Laotiano										

Testo da copiare: Copia

U+0041 LATIN CAPITAL LETTER A

Tabella caratteri

File Visualizza Cerca Vai Aiuto

Times New Roman ▾ Grassetto Corsivo 26 - +

Blocchi Unicode	Tabella caratteri										
Tutti	Dettagli carattere										
Latino base	2	3	4	5	6	7	8	9	:	;	
Supplemento Latino-1	<	=	>	?	@	A	B	C	D	E	
Latino esteso-A	F	G	H	I	J	K	L	M	N	O	
Latino esteso-B	P	Q	R	S	T	U	V	W	X	Y	
Estensioni IPA	Z	[\]	^	_	`	a	b	c	
Lettere che modificano la ...	d	e	f	g	h	i	j	k	l	m	
Segni diacritici combinabili	n	o	p	q	r	s	t	u	v	w	
Greco e copto	x	y	z	{		}	~				
Cirillico											
Supplemento cirillico											
Armeno											
Ebraico											
Arabo											
Siriano											
Supplemento arabo											
Thaana											
N'Ko											
Samaritano											
Mandaico											
Supplemento siriaco											
Arabic Extended-B											
Arabo esteso-A											
Devanagari											
Bengali											

Testo da copiare: Copia

U+0041 LATIN CAPITAL LETTER A

Tabella caratteri

- □

File Visualizza Cerca Vai Aiuto

Times New Roman ▼ Grassetto Corsivo 26 - +

Blocchi Unicode Tabella caratteri Dettagli carattere

Tutti

Latino base

Supplemento Latino-1
Latino esteso-A
Latino esteso-B
Estensioni IPA
Lettere che modificano la ...
Segni diacritici combinabili
Greco e copto
Cirillico
Supplemento cirillico
Armeno
Ebraico
Arabo
Siriano
Supplemento arabo
Thaana
N'Ko
Samaritano
Mandaico
Supplemento siriaco
Arabic Extended-B
Arabo esteso-A
Devanagari
Bengali

U+0041 LATIN CAPITAL LETTER A

Proprietà generali del carattere

Nella categoria Unicode da: 1.1
Categoria Unicode: Lettera, maiuscolo

Varie rappresentazioni utili

UTF-8: 0x41
UTF-16: 0x0041
UTF-8 escaped ottale C: \101
Entità XML decimale: A

Testo da copiare: Copia

Tabella caratteri

File Visualizza Cerca Vai Aiuto

Times New Roman ▾ Grassetto Corsivo 26 – +

Blocchi Unicode

- Devanagari esteso
- Kayah Li
- Rejang
- Hangul Jamo esteso-A
- Javanese
- Birmano esteso-B
- Cham
- Birmano esteso-A
- Tai Viet
- Estensioni meetei mayek
- Etiopia esteso-A
- Latino esteso-E
- Supplemento cherokee
- Meetei Mayek
- Sillabe hangul
- Hangul Jamo esteso-B
- Surrogati alti
- Surrogati per uso privato ...
- Surrogati bassi

Area uso privato

- Ideogrammi compatibilit...
- Forme di presentazione a...
- Forme di presentazione a...
- Selettori di variazioni
- Forme verticali

Tabella caratteri Dettagli carattere



U+E040 <Uso privato>

Proprietà generali del carattere

Nella categoria Unicode da: 1.1
Categoria Unicode: Altro, uso privato

Varie rappresentazioni utili

UTF-8: 0xEE 0x81 0x80
UTF-16: 0xE040

UTF-8 escaped ottale C: \356\201\200
Entità XML decimale: 

Testo da copiare:

[Home](#)[Browse by character](#)[Browse by code chart](#)[Browse by range](#)[Browse by updates](#)[Full code chart](#)[MUFI: The Medieval Unicode Font Initiative](#)[3. MUFI fonts](#)

3. MUFI fonts

This page contains a list of fonts that comply with

- (a) The Unicode Standard
 - (b) the allocation of characters in the Private Use Area of The Unicode Standard, as specified in the MUFI character recommendation
- Please note that the downloading of fonts on this site is entirely at each user's own risk.

To the best of our knowledge, no fonts on this page violate any copyright. Anyone who wishes to make a copyright claim should contact the type designer at the appropriate web sites or any of the MUFI contacts.

Contents

A. Fonts with extensive coverage of MUFI characters

- **Alphabetum** - MUFI v. 1.0 compatible, and a good, but not complete, coverage of MUFI v. 2.0 and v. 3.0.
- **Andron Scriptor** - compatible with the latest version of MUFI at 3 July 2020.
- **Cardo** - MUFI v. 3.0 compatible.
- **Junicode** - MUFI v. 1.0 fully compatible, extensive coverage of MUFI v. 3.0 and later.
- **LeedsUni** - MUFI v. 2.0 compatible.
- **Palemonas MUFI** - MUFI v. 3.0 compatible. All four styles!
- **Titus Cyberbit** - MUFI v. 1.0 compatible.

B. Fonts with limited coverage of MUFI characters

- **Lapidaria Menotec** - contains a selection of Old Norse characters.
- **DejaVu Menotec** - contains a selection of Old Norse characters, and several other code charts.

subsections

- 1. [Alphabetum](#)
- 2. [Andron Scriptor](#)
- 3. [Cardo](#)
- 4. [Junicode](#)
- 5. [LeedsUni](#)
- 6. [Palemonas MUFI](#)
- 7. [Titus Cyberbit](#)
- 8. [Lapidaria Menotec \(for Old Norse\)](#)
- 9. [DejaVu Menotec \(for Old Norse\)](#)

 [log in](#)

Home	0020	0021	0022	0023	0024	0025	0026	0027	0028	0029	002A	002B	002C	002D	002E	002F
	!	"	#	\$	%	&	'	()	*	+	,	-	.	/	
Browse by character	0030	0031	0032	0033	0034	0035	0036	0037	0038	0039	003A	003B	003C	003D	003E	003F
Browse by code chart	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
Browse by range	0040	0041	0042	0043	0044	0045	0046	0047	0048	0049	004A	004B	004C	004D	004E	004F
Browse by updates	@	0041														
Full code chart	P	LATIN CAPITAL LETTER A														
	0060	default font														
	'	d	e	f	g	h	i	j	k	l	m	n	o			
	0070	A	0074	0075	0076	0077	0078	0079	007A	007B	007C	007D	007E	007F		
	P	t	u	v	w	x	y	z	{		}				~	
	00A0	00A4	00A5	00A6	00A7	00A8	00A9	00AA	00AB	00AC	00AD	00AE	00AF			
	00B0	À	À	À	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï		
	o	00B4	00B5	00B6	00B7	00B8	00B9	00BA	00BB	00BC	00BD	00BE	00BF			
	00C0	Entity: — (HTML)	00C4	00C5	00C6	00C7	00C8	00C9	00CA	00CB	00CC	00CD	00CE	00CF		
	À	Ä	Ä	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï			
	Range: BasLat	00D4	00D5	00D6	00D7	00D8	00D9	00DA	00DB	00DC	00DD	00DE	00DF			
	D	00D0	Unicode version: 5.1													
	00E0	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß			
	à	á	ä	å	æ	ç	è	é	ê	ë	í	í	î	ï		
	à	a	a	a	ä	å	æ	ç	è	é	ê	ë	í	í		
	00F0	00F1	00F2	00F3	00F4	00F5	00F6	00F7	00F8	00F9	00FA	00FB	00FC	00FD	00FE	00FF
	ð	ñ	ò	ó	ô	ö	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ
	0100	0101	0102	0103	0104	0105	0106	0107	0108	0109	010A	010B	010C	010D	010E	010F
	Ā	ā	Ā	ă	Ą	ą	Ć	ć	Ę	ę	Ċ	ċ				
	0110	0111	0112	0113	0114	0115	0116	0117	0118	0119	011A	011B	011C	011D	011E	011F
	Đ	đ	Ē	ē	Ę	ě	Ę	é	Ę	ę	Ċ	ċ				
	0120	0121	0122	0123	0124	0125	0126	0127	0128	0129	012A	012B	012C	012D	012E	012F
	Ğ	ğ					h				ī	ī	ī	ī	ī	ī

[Home](#)[Browse by character](#)[Browse by code chart](#)[Browse by range](#)[Browse by updates](#)[Full code chart](#)

Browse by code chart: PUA-36

The Unicode Standard v. 5.1 has no precomposed characters with both acute accent and ogonek. The present subrange contains nine precomposed characters with this combination. The Unicode Standard advises encoding characters of this type as a sequence of the baseline character and the COMBINING OGONEK (0328) and the COMBINING ACUTE ACCENT (0301), and this way of encoding has been given as an alternative for each character below. Note that two of the characters in the list below have duplicate code points in Junicode (v. 0.6.12): E20C and E60C.

 Filter items...

E004

LATIN CAPITAL LETTER A WITH OGONEK AND ACUTE

E099

LATIN CAPITAL LETTER E WITH OGONEK AND ACUTE

E20C

LATIN CAPITAL LETTER O WITH OGONEK AND ACUTE

E257

LATIN CAPITAL LETTER O WITH STROKE AND OGONEK AND ACUTE

E404

LATIN SMALL LETTER A WITH OGONEK AND ACUTE

E499

LATIN SMALL LETTER E WITH OGONEK AND ACUTE

E60C

LATIN SMALL LETTER O WITH OGONEK AND ACUTE

E657

LATIN SMALL LETTER O WITH STROKE AND OGONEK AND ACUTE

E004

LATIN CAPITAL LETTER A WITH OGONEK AND ACUTE

default font

Alt. codepoint combination

Entity: &Aogonacute; (MUFI)

= 0041 + 0328 + 0301

Range: PUA-36**Unicode version:** 5.1**MUFI version:** 3.0**Notes**

= LATIN CAPITAL LETTER A + COMBINING OGONEK + COMBINING ACUTE ACCENT

Encoding special characters

- it might be necessary:
 - for paleographic and/or philological analysis
 - to allow precise textual research
 - to be able to accurately reproduce the original
- what needs to be marked:
 - not all "special" characters and their variations are significant
 - you should clearly decide which characters will be marked
- several methods available:
 - TEI XML + Unicode characters + Unicode fonts
 - use of entities and intelligent markup (TEI P4, generic XML)
 - use of the TEI gaiji module (TEI P5)

The TEI *gaiji* module elements

- allows to encode special characters in a **<charDecl>** element
 - part of the **<encodingDesc>** section of the **<teiHeader>**
- which includes three possible sub-elements:
 - **<desc>** a short description (optional)
 - **<char>** a special or non standard character
 - **<glyph>** a special or non standard glyph
- a new element **<g>** is available:
 - to be used in the main text
 - **@ref** attribute to link to **<char>** or **<glyph>** elements
 - may be an empty element or hold a single character

The <char> and <glyph> sub-elements

- <unicodeProp> and <localProp>: properties of a character or glyph as described in the Unicode standard or defined by the encoder

```
<unicodeProp name="Name" value="LATIN ENLARGED LETTER SMALL A"/>
<localProp name="Name" value="LATIN ENLARGED LETTER SMALL A WITH ACUTE"/>
```

- <mapping>: character mapping, one or more characters

```
<char xml:id="aeligacute">
  <unicodeProp name="Name" value="LATIN SMALL LETTER AE WITH ACUTE"/>
  <mapping type="codepoint">U+01FD</mapping>
  <mapping type="diplomatic">é</mapping>
  <mapping type="normalized">æ</mapping>
</char>
```

The <charDecl>

Use of <g> in the text <body>



<body>

```
<div n="DOTR" subtype="edition_text" type="verse" xml:id="DOTR">

<!-- ##### folio 104v ##### -->
<pb n="104v" xml:id="VB_fol_104v"/>

<l n="1"><lb facs="#VB_line_104v_07" n="7" xml:id="VB_lb_104v_07"/><hi rend="init3.1"><g ref="#Hunc"/></hi><hi
rend="cap">W</hi>æt ic <g ref="#slong"/>wefna c<g ref="#ydot"/><g ref="#sins"/>t secgan wylle</l>
<l n="2"><choice>
  <sic>hæt</sic>
  <corr resp="Grein">hwæt</corr>
</choice>
<choice>
  <orig>mege mætte</orig>
  <reg>me gemætte</reg>
</choice><lb facs="#VB_line_104v_08" n="8" xml:id="VB_lb_104v_08"/>to midre nihte</l>
<l n="3">syðpan <choice>
  <orig>reord b<g ref="#eenl"/>r<g ref="#eenl"/>nd</orig>
  <reg>reordberend</reg>
</choice> reste wunedon<orig><pc type="metrical">. </pc></orig></l>
<l n="4"><lb facs="#VB_line_104v_09" n="9" xml:id="VB_lb_104v_09"/>þuhte me þæt ic <choice>
  <orig>ge <g ref="#sins"/>awe</orig>
  <reg>gesawe</reg>
</choice>
<g ref="#sins"/>yllcre treow</l>
<l n="5"><choice>
  <orig>onlyft</orig>
  <reg>on lyft</reg>
</choice><lb facs="#VB_line_104v_10" n="10" xml:id="VB_lb_104v_10"/>lædan leohte <choice>
  <orig>be wunden</orig>
  <reg>bewunden</reg>
</choice></l>
```

Diplomatic level

6
7 **þ**Wæt • ic swefna cýr secgan wylle hæt mege mætte
8 to midre nihte syðþan reord berend reste wunedon. *I*
9 þuhte me þæt ic ge ȝawe syllicre treow onlyft
10 lædan leohte be wunden beama beorhtoþt eall þæt
11 beacen wæſ be gotten mid golde gimmaſ stodon fægere

Interpretative level

6
7 HWæt • ic swefna cyst secgan wylle hwæt me gemætte
8 to midre nihte syðþan reordberend reste wunedon. *I*
9 þuhte me þæt ic gesawe syllicre treow on lyft
10 lædan leohte bewunden beama beorhtost eall þæt
11 beacen wæſ begoten mid golde gimmas stodon fægere



Edition levels

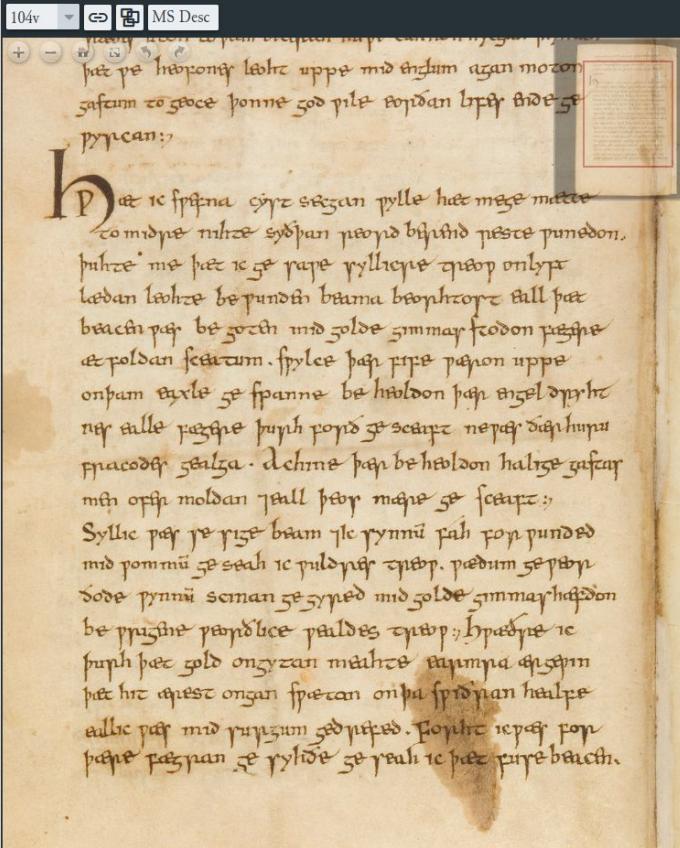


Edition levels

- you can take advantage of TEI encoding to obtain different levels of edition for the same text (single source document):
 - scribal correction elements: and <add>
 - editorial intervention elements: <sic> and <corr>, <orig> and <reg>, <abbr> and <expan>, <supplied>
 - the different <mapping>s provided for special characters
 - other elements: <hi>, <pc>, <damage> etc.
- using EVT it is then possible to automatically generate the different levels of text edition

```
<l n="13"><lb facs="#VB_line_104v_17" n="17" xml:id="VB_lb_104v_17"/><hi rend="cap">S</hi>yllic wæg ref="#sins"/>
<g ref="#sins"/>e <seg type="kenning"><choice>
  <orig><g ref="#sins"/>ige beam</orig>
  <reg>sigebeam</reg>
</choice></seg>
<choice>
  <orig><abbr type="brev"><am><g ref="#et"/></am></abbr><g ref="#inodotenl"/>c</orig>
  <reg><expan>and</ex></expan> ic</reg>
</choice>
<choice>
  <abbr><g ref="#sins"/>ynn<am><g ref="#umacr"/></am></abbr>
  <expan>synnu<ex>m</ex></expan>
</choice> f<g ref="#aacute"/>h</l>
<l n="14"><choice>
  <orig>for wounded</orig>
  <reg>forwunded</reg>
</choice><lb facs="#VB_line_104v_18" n="18" xml:id="VB_lb_104v_18"/><mid <choice>
  <abbr>womm<am><g ref="#umacr"/></am></abbr>
  <expan>wommu<ex>m</ex></expan>
</choice>
<choice>
  <orig>ge seah</orig>
  <reg>geseah</reg>
</choice> ic wuldre<g ref="#sins"/> treow<orig><pc type="metrical">. </pc></orig></l>
```

104v



The Dream of the Rood 104v Diplomatic Info ⓘ

1
2
3
4
5
6

7 **D**Wat • ic swefna cijt secgan wylle hæt mege mætte
8 to midre nihte syðjan reord berend reste wunedon. 7
9 þuhte me hæt ic ge rawe syllcre treow onlyf
10 lædan leahite be wunden beamma beorhtoſt call hæt
11 beacan wæt be gotten mid golde gimmer stodon fægere
12 et fordan sceatum . sywleas hæt fife wæron uppe
13 on þam eadle ge spanne be healdon hæt engel dryht
14 ne calle fæger hƿurh forð ge sceaf newaſ ðærhuru
15 fræcder gealga . aelhine hæt be healdon halige gaſtar
16 men ofer moldan **ȝeall** heor mære ge sceaf .
17 Syllic wæt je rige beam **ȝic** **wynnū** fāh **for** wunded
18 mid **wommu** ge seah ic wuldref treow **forwundē** jewor
19 dode **wynnū** scinan ge gyred mid golde gimmer haefdon
20 be wrigene weordlice wældes treow ; hwaðre ic
21 hƿurh hæt gold ongytan mæhte carmra ærgewin
22 hæt hit arest ongan swætan on þa swiðran healfse
23 eallic wæt mid **wægum** gedrefed . Forht ic wæt for
24 hære fægran ge ƿyhðe ge yeah ic hæt fure beacen .

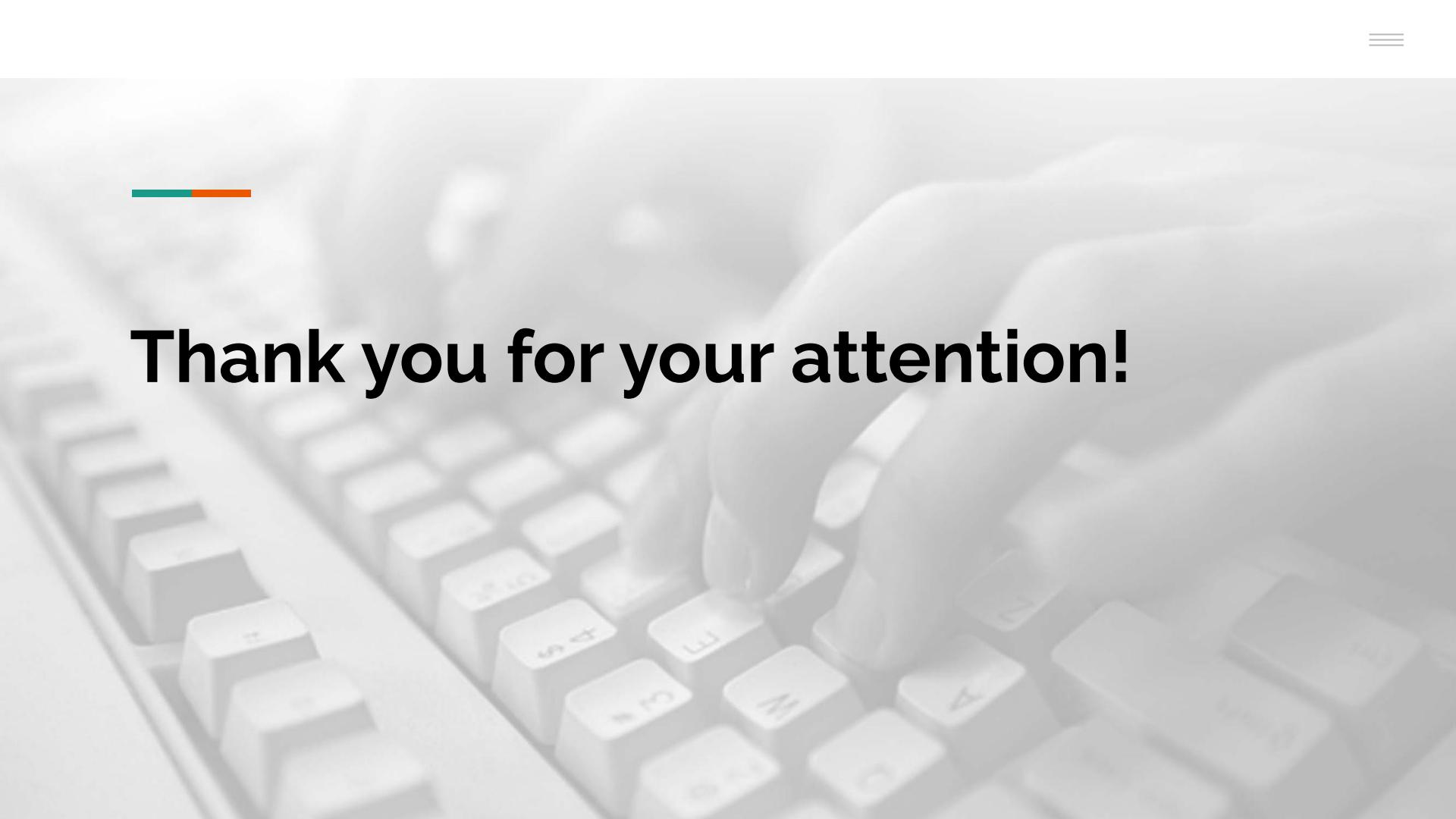
Nessuna Selezione Search Create index A



Other methods

- the Menota schemas (https://www.menota.org/EN_forside.xhtml) allow three edition levels

Facsimile	Transcriptions	Renderings
	<me:fac>le&gscap;ıa</me:fac> <me:dipl>le&gscap;ia</me:dipl> <me:norm>leggja</me:norm>	legja legja leggja
	<me:fac>&slong;ıðan</me:fac> <me:dipl>&slong;iðan</me:dipl> <me:norm>síðan</me:norm>	síðan síðan síðan
	<me:fac>gı&ocurl;&fins;</me:fac> <me:dipl>gi&ocurl;f</me:dipl> <me:norm>giðf</me:norm>	giðf giðf giðf



— ■

Thank you for your attention!