

The Values of French Language and Literature in the European Middle Ages
***The Histoire ancienne jusqu'à César: A Digital Edition* (3.0)**

TEI-XML – Conversion

Introduction

This document outlines the employment of TEI-XML in *The Histoire ancienne jusqu'à César: A Digital Edition*. It does not include the more standard uses of TEI-XML. The document also details the conversion of the encoding to render the semi-diplomatic and interpretive versions of the edition in the Text Viewer (<https://tvof.ac.uk/textviewer/>).

Contents

- 1. Text Block**
- 2. Medieval Punctuation**
- 3. Transcription: Additions/Deletions/Modifications/Annotations**
- 4. Discursive/Linguistic Features**
- 5. TVOF Editorial Interventions**
- 6. Editorial Notes**
- 7. Images of the Manuscript**
- 8. Text Viewer: Key**

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
1. Text Block			
<head type="rubric">XXX</head>	Rubricated heading.	Text rendered in red, larger than paragraph body, preceded by the paragraph number and full stop.	Same as semi-dip.
'Paragraph' body (prose): <ab> <seg type="1" xml:id="edfr20125_00001_01">XXX</seg> <seg type="1" xml:id="edfr20125_00001_02">XXX</seg> <seg type="1" xml:id="edfr20125_00001_03">XXX</seg> </ab> For Fr20125, we use @rend="NR" when the first letter of the <seg> is not highlighted with a stroke of red ink: <seg type="1" rend="NR" xml:id="edfr20125_00xxx_01"> Possible <seg> @types for prose: 1/2/3/4/5/-a (see appendix 1 below)	'Paragraph' in prose. The delimitation of a <seg> follows the medieval punctuation, i.e. the use of a punctuation mark and an uppercase letter by the scribe (highlighted with a stroke of red ink in Fr20125). 	Text in each <seg> preceded by segment number (XXXXX_01).	Same as semi-dip, but render all of the text in black.
'Paragraph' body (verse): <ab> <seg type="6" xml:id="edfr20125_00001_01"> <lg type="octo_coup"> <lg type="lineated"> <l n="001">XXX</l> <l n="002">XXX</l> </lg></lg></seg></ab> <seg> @type for verse: 6	'Paragraph' in verse. All verse passages are encased in a single <seg> and a line group (<lg>). The individual lines are numbered consecutively. The first <lg> indicates the form: <lg type="octo_coup">	Line numbers: every 4 lines; superscript at the end of the line.	Line numbers = same as semi-dip.

We note the mise-en-page using: <lg type="lineated"> or <lg type="cont">	= octosyllabic couplet The second <lg> indicates the mise-en-page: <lg type="lineated"> = copied line by line <lg type="cont"> = written continuously	<lg type="lineated">: rendered line by line and line group highlighted in purple <lg type="cont">: rendered continuously and line group highlighted in purple	<lg type="lineated"> + <lg type="cont">: rendered line by line and line group highlighted in purple
'Paragraph' body (prose + verse): <ab> <seg type="1" xml:id="edfr20125_00001_01">XXX</seg> <seg type="1" xml:id="edfr20125_00001_02">XXX</seg> <seg type="6" xml:id="edfr20125_00001_03"> <lg type="octo_coup"> <lg type="lineated"> <l n="001">XXX</l> <l n="002">XXX</l> </lg></lg></seg></ab>	'Paragraph' in prose and verse. Verse passage placed in a <seg> and <lg> (see description above).	Prose Precede text in each <seg> by segment number (XXXXX_01). Verse Line numbers: every 4 lines; superscript at the end of the line. <lg type="lineated">: render line by line and highlight line group in purple <lg type="cont">: render continuously and highlight line group in purple.	Prose Same as semi-dip. Verse <lg type="lineated"> + <lg type="cont">: render line by line and highlight line group in purple
Prose + verse in a single <seg> <seg type="1" subtype="6" xml:id="edfr20125_00007_03">XXXX<pc rend="1"/> <lg type="octo_coup"> <lg type="cont"> <l n="001">XXX</l> <l n="002">XXX</l> </lg></lg></seg> Use @type as for prose (see above) and add @subtype= 6.	See above.	As above.	As above.

<code><seg type="explicit"></code> Example: Roy §577 <code><seg type="explicit"</code> <code>xml:id="edRoyal20D1_00577_explicit_Orient_II"></code>	Paratextual content marking the end of a narrative section.	Add a line break before text enclosed in <code><seg type="explicit"></code>	Same as semi-dip.
<code><pb></code> + <code><cb></code> Examples: <code><pb n="29v"/><cb n="a"/></code> <code><cb n="b"/></code>	Page/column breaks	Render in square brackets. Where the milestone divides a textual string add dashes at either side, e.g. - [XXX]-	Same as semi-dip.

Appendix 1

@types of `<seg>` (prose)

- `<seg type="1">` = 1 clause, Ø dependent (embedded or other) clause
- `<seg type="2">` more than 1 root clause, but Ø dependent (embedded or other) clause
- `<seg type="3">` 1 root clause + 1 or more dependent (embedded or other) clause(s)
- `<seg type="4">` 2 root clauses + dependent (embedded or other) clauses
- `<seg type="5">` more than 2 root clauses + dependent (embedded or other) clauses
- -a is used when there is syntactical continuity linking the `<seg>` to the previous one

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
2. Medieval punctuation			
<code><pc rend="1"/></code>	punctus 	midpoint (U+00B7)	Not rendered in interpretive text.
<code><pc rend="2"/></code>	punctus elevatus 	inverted-semicolon (U+061B)	Not rendered in interpretive text.
<code><pc rend="3"/></code>	punctus interrogativus 	question-mark (U+003F)	Not rendered in interpretive text.
<code><pc rend="4"/></code>	line break in prose	'[/]' = forward slash (U+002F) in square brackets	Not rendered in interpretive text.
<code><pc rend="5"/></code>	absence of punctuation	[blank]	Not rendered in interpretive text.
<code><pc rend="6"/></code> Example: <code><seg type="1" xml:id="edRoyal20D1_00610_26"><pc rend="6"/>Quant...</seg></code>	[blank space+]//: 'paragraphing' system; possibly used to mark the position of a pied-de- mouche. 	// = double forward slash (U+002F) Tool tip explains: "probablement pour designer la position d'un pied-de-mouche"	Paragraph break + indent
<code><pc rend="7"/></code> Example: <code><seg type="1" xml:id="edRoyal20D1_00116_04"><pc rend="7"/>La seconde...</seg></code>	Pilcrow 	Pilcrow (U+00B6)	Not rendered in the interpretive text.

Use of red highlight and red ink			
<hi rend="RH">	A single letter highlighted with a stroke of red ink, which does not occur at the start of a <seg>.	Encased letter rendered in black.	Encased letter rendered in black.
<hi rend="R">	Content written in red ink, which is not a rubricated heading (e.g. rubricated explicit).	Encased content rendered in red.	Same as semi-dip.

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
3. Transcription: Additions/Deletions/Modifications/Annotations			
<add> @place: where the item is placed. Values: 1] below ; 2] above ; 3] inline; 4] right_margin; 5] left_margin; 6] central_margin; 7] top_margin; 8] bottom_margin; 8] overleaf. @hand: indicates the hand responsible. Values: 1. S = "scribe" 2. E = "medieval editor" 3. R = "rubricator" 4. CE = "cursive editor" 5. U = "unknown" @cert has been used to indicate level of certainty. Values: 1] high; 2] low; 3] medium; 4] unknown. Examples: Fr20125 § 935.5 <persName><add place="right_margin" hand="E" cert="medium">rufus</add> minusius</persName> Fr20125 § 946.4 Et <persName>scipio cornelius</persName> <add place="bottom_margin" hand="E">u,ers les mons por sau,oir se <persName>hanibal</persName> se u,oudroit cele part t[ra]ire<pc rend="1"/> </add>	Additions by a scribe, rubricator, medieval editor or later reader	The content is enclosed in angle brackets. A tooltip indicates the text is the result of an addition and includes the information in @hand in French, see below: 1. S = "scribe" = copiste 2. E = "medieval editor" = rédacteur médiéval 3. R = "rubricator" = rubricateur 4. CE = "cursive editor" = rédacteur médiéval en cursive 5. U = "unknown" = inconnue If there is no indication of @hand, it is as yet unspecified.	Text enclosed in <add> is included in the main text body without any angle brackets.

 @hand: indicates the hand responsible. Values: 1. S = “scribe” 2. E = “medieval editor” 3. R = “rubricator” 4. U = “unknown” @rend: how the deletion was rendered or presented in the manuscript. Values: 1. strikethrough 2. multiple_strikethrough 3. expunctuation 4. scraped	Deletions by a scribe, rubricator, medieval editor or later reader.	Strikethrough the deleted text	Deleted content is not included in the text body.
<mod> + + <add> Example: Roy20 § 579.18 <mod hand="E"> <add>g</add>i</mod>ent	Modifications to letter forms or textual strings by medieval scribes/editors.	The content in <add> is presented in angle brackets (the same angle brackets as those used for <add> alone, see above). A tooltip indicates that it is a “modification” and the content in is included after the words “contenu original:”. The content in @hand in <mod> is rendered in French as below: 1. S = “scribe” = “copiste” 2. E = “medieval editor” = “rédacteur médiéval” 3. U = “unknown” = “inconnu” If there is no indication of @hand, it is as yet unspecified.	Only content in <add> is included in the main text body, without angle brackets (as <add> above).
<note type=“gloss”> @place: where the item is placed.	Marginal annotations by a	A speech icon indicates that there is a marginal annotation.	The medieval marginal notes are not included in the interpretive text.

Values:1] below ; 2] above ; 3] inline; 4] right_margin; 5] left_margin; 6] central_margin; 7] top_margin; 8] bottom_margin; 9] overleaf. @resp indicates the hand responsible: 1. S = “scribe” 2. E = “medieval editor” 3. R = “rubricator” 4. CE = “cursive editor” = rédacteur médiéval en cursive 5. LH = “later hand” (responsible for marginal glosses in Catalan in f.fr. 20125) 6. LH2 = “later hand 2” 7. D = “duplicator” 8. U = “unknown” @cert has been used to indicate level of certainty. Values: 1] high; 2] low; 3] medium; 4] unknown. Example: Fr20125 § 2.14 <note type="gloss" place="right_margin" resp="LH">no[ta] d[e] p[ri]mer le[n]gatge</note>	reader of the manuscript.	The tooltip includes the text transcribed in <note type="gloss"> and the information in @resp in French, see below: 1. S = “scribe” = “copiste” 2. E = “medieval editor” = “rédacteur médiéval” 3. R = “rubricator” = “rubricateur” 4. CE = “cursive editor” = rédacteur médiéval en cursive 5. LH = “later hand” = “annotateur catalan?” 6. LH2 = “later hand 2” = “annotateur X” 7. D = “duplicator” = “annotateur D” 8. U = “unknown” = “inconnu” If there is no indication of resp, it is as yet unspecified. Do not include the information in @cert, if used. In the print screen viewing option, the annotation is displayed in the same manner as the editorial notes that accompany the interpretive text, as below: ‘Note de lecteur médiéval ([hand, e.g. ‘annotateur D’]) : [content of annotation]’	
<unclear> @reason: succinct account of why the passage is materially hard to transcribe. Values: abbreviation; illegible @agent: cause of damage if identifiable.	Text is unclear on the manuscript page.	Question mark icon + tooltip “texte illisible”	Content in <unclear> is not included in the interpretive text.

Values: rubbing, mildew, smoke			
@cert has been used to indicate level of certainty. Values: 1] high; 2] low; 3] medium; 4] unknown.			

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
4. Discursive/Linguistic Features			
<gloss>	Explanation of a word/expression.	Do not highlight	Same as semi-dip.
<quote source="X">	Quotation from an identified source.	Do not highlight	Same as semi-dip.
<foreign xml:lang="X">	Words transcribed in languages other than French.	Do not highlight	Same as semi-dip.
<q type="NI">	Narratorial interventions.	Do not highlight	Same as semi-dip.
<said direct="true/false">	Direct/indirect speech.	Do not highlight	Same as semi-dip.
<title>	Title of another work.	Do not highlight	Same as semi-dip.
<rs type="author/metatext/metaling/US">	Used to encode references to: - the author of the HA; - language/a language; - the diegesis or sources of the HA; - unspecified sources.	Do not highlight	Same as semi-dip.

<name> or <pers-/geog-/placeName> For references to authors, use: <persName type="author"> Uses of <name> @type: <name type="people">romains</name> <name type="building">temple</name> <name type="order">cisteaus</name>	Names, places, nationalities, etc.	Do not differentiate from main text.	The first letter of each name is capitalised.
---	---	---	--

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
5. TVOF Editorial Interventions			
<choice> + <sic> + <corr> Example: Fr20125 § 596.12 <choice> <sic>quil aseur quil aseur</sic> <corr>qu<choice><orig /><reg>'</reg></choice>il aseur</corr></choice>	Used to encode the manuscript reading (<sic>) vs the text “corrected” by the TVOF editor (<corr>) (e.g. correction of duplications, or emendations for the sake of clarity).	Only content in <sic> is displayed.	Content in <corr> is displayed in angle brackets. Content in <sic> is presented in a tooltip preceded by <i>ms.</i> [=ms reading].
<supplied> @resp=TVOF Example: Fr20125 § 893.6 p<supplied resp="TVOF"> </supplied>us	Additions by the modern editor to render the interpretive text more comprehensible.	Not rendered in the semi-diplomatic text.	Content in <supplied> rendered in angle brackets.
<choice>+<seg type="semi-dip">+<seg type="crit"> Example: Fr20125 § 590.2 <choice><seg type="semi-dip">dela</seg><seg type="crit">de la</seg></choice>	Used to encode the normalisation of spacing and letter forms, and the addition of modern punctuation, etc., in the interpretive text.		

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
6. Editorial Notes			
Source Notes Example: <pre><div subtype="sourceNotes" type="notes" xml:id="edRoyal20D1_sourceNotes"> <div subtype="source" type="note" xml:id="edRoyal20D1_00527_peach"> <head type="noteLabel"/> <p xml:id="edRoyal20D1_00527_peach_a"><hi rend="i">Aen.</hi>, I, 88-222, 305-313, 418-420</p> </div> </div></pre>	Notes on sources added by the TVOF editor.	Not rendered in the semi-diplomatic text.	Purple speech bubble icon. Content in <p> included in the note pop-up.
General Notes Example: <pre><div type="notes" subtype="genNotes" xml:id="edRoyal20D1_genNotes"> <div type="note" subtype="gen" xml:id="edRoyal20D1_00525_02_cairo"> <head type="noteLabel"/> <p xml:id="edRoyal20D1_00525_02_cairo_a">XXX</p> </div> </div></pre>	Notes on any other subject added by the TVOF editor, including variant readings.	Not rendered in the semi-diplomatic text.	Green speech bubble icon. Content in <p> in the note pop-up.
<anchor> Positioned inside the relevant <head> (sources of a paragraph), <seg> or <l>.	Used to indicate the position of the note.		

Example: Fr20125 § 590 <anchor type="sourceNote" corresp="#edfr20125_0590_peach"/>			
--	--	--	--

Encoding	Description	Rendering: Semi-diplomatic	Rendering: Interpretive
7. Images of the Manuscript			
<figure> Can be placed within a <seg>, or anywhere in a <div>. Example: Fr20125 <figure><graphic url="0322_FRANCAIS-20125_folio-156v.jp2"/> <head>Suicide de Didon</head> <p>Source : Bibliothèque nationale de France</p></figure>	A reference to a full-page image of the manuscript featuring illustration(s).	An inline camera icon indicates the presence of an image. If the user clicks the image, a modal pops up with a zoomable image of the manuscript page. The content in <head> appears above the image. The content in <p> appears below the image.	Same as semi-dip.

8. Text Viewer Key

Semi-diplomatic	Interpretive
[999ra] Feuillet, recto/verso, colonne	[999ra] Feuillet, recto/verso, colonne
[abc] Résolutions d'abréviations dans le ms	<abc> Texte ajouté ou corrigé par l'éditeur moderne
<abc> Texte ajouté à la transcription par une main médiévale	abc {purple background} Texte en vers
abc Texte effacé dans le ms	{purple note icon} Note sur les sources
[...] Lacune textuelle/Texte gratté	{green/turquoise note icon} Note générale
abc {purple background} Texte en vers	