

TEI BY EXAMPLE

<**TBE:eg**>
TEI
By
Example
</**TBE:eg**>

MODULE 2: THE TEI HEADER

Ron Van den Branden

Edward Vanhoutte

Melissa Terras

Centre for Scholarly Editing and Document Studies (CTB) , Royal
Academy of Dutch Language and Literature, Belgium, Gent, 9 July 2010

Last updated September 2020

Licensed under a Creative Commons Attribution ShareAlike 3.0 License

TABLE OF CONTENTS

4. The Header of a Complex Text.....	1
--------------------------------------	---

4. The Header of a Complex Text

Before we end, let's go back to where we left you: the library, in front of the library catalogue or computer screen. Prepared for the possibility that this copy of the book may be in loan, you find another reference to *The Wild Ass's Skin* in the record of La Comédie Humaine (look for “505 8 0 |gPhilosophic and analytic studies: v. 41. The|tmagic skin”; *The Magic Skin* is an alternative title for the English translation):

Holdings	Comments	Librarian's View	Copyright Status
000	03482cam a2200265 a 4500		
001	3724946		
005	20060802115229.0		
008	030626m18951900enkcf 000 1 eng d		
019 1	a24758250		
040	aNU beng cNU dANL		
041 1	aeng hfre		
100 1	aBalzac, Honoré de, d1799-1850.		
240 1 0	aComédie humaine. lEnglish		
245 1 4	aThe human comedy / cby Honoré de Balzac.		
260	aLondon : bCaxton, c1895-1900.		
300	a53 v. : bports. : c22 cm.		
500	aEach volume has separate t-p.		
505 0 0	gParisian life: v. 1-2. The tsplendors and miseries of courtesans g; The tway that girls love -- tHow much love costs old men -- gThe tend of bad roads -- gThe tlast incarnation of Vautrin -- gv. 3-5 tPoor relations tCousin Bette -- tCousin Pons -- gv. 6 tHistory of the thirteen tFerragus, Chief of the Devorants -- gLa tDuchesse de Langeais -- gv. 7 tHistory of the grandeur and downfall of César Birotteau -- gv. 8. The tcivil service -- tGaudissart II -- gA tprince of Bohemia -- gv. 9. The thouse of Nucingen -- gThe tsecrets of La Princesse de Cadignan -- tSarrasine -- tFacino Cane -- gA tman of business -- gThe tinvoluntary comedians -- gv. 10-11. The tpetty bourgeois.		
505 8 0	gPrivate life: v. 12. The thouse of the Cat and Racket -- gThe tdance at Sceaux -- gThe tpurse -- gThe tvendetta -- gv. 13 tMemoirs of two young wives -- gv. 14 tModeste Mignon -- gv. 15. A tstart in life -- tMadame Firmiani -- gThe tmessage -- gThe tatheist's mass -- gv. 16 tAlbert Savarus -- gA tdaughter of Eve -- gv. 17. The tmarriage contract -- gLa tgrenadiere -- tGobseck -- gv. 18. A twoman of thirty -- gThe tdeserted mistress -- v. 19. tOld Goriot -- v. 20 tHonorine -- tColonel Chabert -- gThe tinterdiction -- gv. 21 tBeatrix -- gv. 22. A tdouble family -- gThe tpeace of the household -- gA tstudy of woman -- tAnother study of woman -- gThe tpretended mistress.		
505 8 0	gProvincial life: v. 23 tUrsule Mirouet -- gv. 24-25. The tcelibates tPierrette -- gThe tCuré de Tours -- gLa tRabouilleuse -- gv. 26 tEugénie Grandet -- gv. 27. The tily of the valley -- gv. 28. The tillustrious Gaudissart -- gThe tmuse of the department -- gv. 29. The told maid -- gThe tcabinet of antiquities -- gv. 30-32. tLost illusions g:The ttwo poets -- gA tprovincial great man in Paris -- gThe ttrials of an inventor.		
505 8 0	gMilitary and political life: v. 33. The tChouans -- gv. 34. A tpassion in the desert -- gAn tepisode under the reign of terror -- gA tdark affair -- gv. 35 tZ. Marcas -- gThe tother side of contemporaneous history gfirst episode tMadame de la Chanterie g; second episode: The tnovice -- gv. 36-37. The tdeputy from Arcis.		
505 8 0	gCountry life: v. 38. The tcountry Doctor -- gv. 39. The tvillage Curé -- gv. 40. The tpeasants.		
505 8 0	gPhilosophic and analytic studies: v. 41. The tmagic skin -- gv. 42. tJesus Christ in Flanders -- gv. 43. The tunknown masterpiece -- gThe tMaranas -- gA tseashore drama -- gThe tred inn -- tMaster Cornelius -- gv. 44 tGambara -- tMassimilla Doni -- gThe taccursed child -- gv. 46. The tquest of the absolute -- gv. 46 tAdieu -- gThe tconscript -- gThe texecutioner -- gThe texiles -- tLouis Lambert -- gv. 47 tCatherine de' Medici -- gv. 48-49 tPhysiology of marriage -- tPetty worries of conjugal life.		
505 8 0	gv. 50-51. The trepository of La comédie humaine of Honoré de Balzac -- gv.52-53. Les tcontes drolatiques.		
600 1 0	aBalzac, Honoré de, d1799-1850.		

Figure 4. A [library catalogue record](#) for a complex work.

Now that's a record! If you thought the truckload of possibilities for the description of electronic texts in the TEI header set your head spinning already, imagine what an electronic edition of *La Comédie Humaine* might look like! Code 300 tells us that it has no less than 53 volumes, with different titles per volume.

One way of encoding this majestic work as a whole would be to treat *La Comédie Humaine* as a kind of “supertext” containing all different works. This can be done in TEI by treating the whole as a <`teiCorpus`>, containing each separate work in its own <`TEI`> text. As each of these <`TEI`> texts needs its own TEI header, you can imagine the amount of meta-information, much of which will have to be repeated. This can be avoided by placing the common meta-information in the <`teiHeader`> element of the <`teiCorpus`> element, while retaining all work-specific meta-information in the TEI header section of the respective <`TEI`> text. This mechanism allows you to be maximally expressive in the description of all texts in a TEI corpus, *and* maximally efficient in the reduction of common information in the individual TEI headers.

The following example gives an impression of what a TEI header for an electronic edition of *La Comédie Humaine* might look like:

```
<teiCorpus xmlns="http://www.tei-c.org/ns/1.0">
    <!-- general TEI header information or the entire corpus -->
    <teiHeader>
        <fileDesc>
            <titleStmt>
                <title>La Comédie Humaine</title>
                <author xml:id="HdB">Honoré de Balzac</author>
                <editor role="editor" xml:id="TBEcrew">The TBE crew</editor>
                <respStmt>
                    <name xml:id="RvdB">Ron Van den Branden</name>
                    <resp>transcription</resp>
                    <resp>annotation</resp>
                </respStmt>
                <sponsor>Association for Literary and Linguistic Computing (ALLC)</sponsor>
                <sponsor>Centre for Data, Culture and Society, University of Edinburgh,
                    UK</sponsor>
                <sponsor>Centre for Computing in the Humanities (CCH) - King's College
                    London</sponsor>
                <sponsor>University College London (UCL)</sponsor>
            <funder>
                <address>
                    <addrLine>Centre for Scholarly Editing and Document Studies (CTB)</addrLine>
                    <addrLine>Royal Academy of Dutch Language and Literature</addrLine>
                    <addrLine>Koningstraat 18</addrLine>
                </address>
            </funder>
        </fileDesc>
    </teiHeader>
</teiCorpus>
```

```
<addrLine>9000 Gent</addrLine>
<addrLine>Belgium</addrLine>
</address>
<email>ctb@kantl.be</email>
</funder>
<principal xml:id="EV">Edward Vanhoutte</principal>
<principal xml:id="MT">Melissa Terras</principal>
</titleStmt>
<extent>0.5 Gb</extent>
<publicationStmt>
<publisher>Centre for Scholarly Editing and Document Studies (CTB)</publisher>
<distributor>Centre for Computing in the Humanities (CCH) - King's College
London</distributor>
<pubPlace>Gent</pubPlace>
<address>
<name type="institution">Centre for Scholarly Editing and Document Studies (C
TB)</name>
<name type="institution">Royal Academy of Dutch Language and Literature</name>
<street>Koningsstraat 18</street>
<postCode>9000</postCode>
<name type="city">Gent</name>
<name type="country">Belgium</name>
</address>
<idno type="ISBN">0-00-00000-9</idno>
<availability status="free">
<licence>Published under a <ref target="http://creativecommons.org/licenses/by
-sa/3.0/">Creative Commons Attribution ShareAlike 3.0 License</ref>.</licence>
</availability>
<date when="2110-01-01">1 January 2110</date>
</publicationStmt>
<seriesStmt>
<title>The TBE collection: sample texts encoded with TEI.</title>
<respStmt>
<name>Edward Vanhoutte</name>
<resp>compiler</resp>
</respStmt>
<idno type="ISSN">0000-0001</idno>
</seriesStmt>
```

```
<notesStmt>
    <note>OCR scanning done at KANTL, Ghent.</note>
</notesStmt>
<sourceDesc>
    <bibl><title>La Comédie Humaine</title> by <author>Honoré
        de Balzac</author>. <pubPlace>London</pubPlace>
        : <publisher>Caxton</publisher>, <date from="1895" to="1900">1895
        -1900</date>. <extent>53 v.</extent>.</bibl>
</sourceDesc>
</fileDesc>
<encodingDesc>
    <editorialDecl>
        <correction method="markup">
            <p>Apparent errors have been corrected using the <gi>sic</gi>> <gi>corr</gi>
            elements, wrapped in a <gi>choice</gi> element.</p>
        </correction>
        <normalization method="markup" source="http://www.oed.com/">
            <p>Spelling has been modernised using the <gi>orig</gi> / <gi>reg</gi>
            elements, wrapped in a <gi>choice</gi> element.</p>
        </normalization>
        <quotation marks="all">
            <p>Diplomatic transcription, all original quotation marks have been retained
            and normalised to double quotation marks.</p>
        </quotation>
        <hyphenation eol="none">
            <p>End-of-line hyphenation has been removed. All other hyphenation has been
            retained.</p>
        </hyphenation>
        <interpretation>
            <p>Thematic analysis added, studying the main motifs.</p>
            <p>Names and dates are marked.</p>
        </interpretation>
    </editorialDecl>
    <projectDesc>
        <p>Text encoded for <soCalled>The TBE collection: sample texts encoded with T
        EI</soCalled>, aiming at providing a collection of prime exemplar TEI encoded
        materials.</p>
    </projectDesc>
```

```
<refsDecl>
<p>The paragraphs in the text are numbered with the <att>n</att> attribute.
Each number consists of four digits; numbering is consecutive throughout the
book. For example: <val>0203</val> numbers the 203th paragraph throughout the
book.</p>
<p>Each chapter is identified with a formal identification code inside
the <att>xml:id</att> attribute. Chapters are numbered using arabic numerals.
The codes are composed by concatenating the identification codes for all
ancestor text divisions down to the chapter level, with the dot as separation
marker. For example: <val>I.2.3</val> identifies the third chapter of the
second book of the first volume.</p>
</refsDecl>
<classDecl>
<taxonomy xml:id="DDC">
<bibl>
<title>Dewey Decimal Classification</title>
<edition>Abridged Edition 14</edition>
<ptr target="http://www.oclc.org/dewey/versions/abridgededition14/default
.htm"/>
</bibl>
</taxonomy>
<taxonomy xml:id="lcsh">
<bibl>
<title>Library of Congress Subject Headings</title>
</bibl>
</taxonomy>
</classDecl>
<metDecl>
<metSym value="-">long syllable</metSym>
<metSym value="u">short syllable</metSym>
<metSym value="| ">foot boundary</metSym>
<metSym value="/">line boundary</metSym>
</metDecl>
<variantEncoding method="parallel-segmentation" location="internal"/>
</encodingDesc>
<profileDesc>
<langUsage>
<language ident="en" usage="98">English</language>
```

```
<language ident="fr" usage="1">French</language>
<language ident="ar" usage="1">Arabic</language>
</langUsage>
</profileDesc>
<revisionDesc>
<change when="2109-12-20" who="#RvdB">final proofing</change>
<change when="2009-03-08" who="#MT">addition of thematic analysis</change>
<change when="2009-03-08" who="#RvdB">addition of explanatory notes</change>
<change when="2008-10-25" who="#Rvdb">spell check</change>
<change when="2008-08-25" who="#Rvdb">addition of phrase level markup</change>
<change when="2008-08-20" who="#Rvdb">file creation</change>
</revisionDesc>
</teiHeader>
<!-- ... -->
<TEI>
<!-- specific TEI header information for the distinct TEI documents --&gt;
&lt;teiHeader&gt;
&lt;fileDesc&gt;
&lt;titleStmt&gt;
&lt;title&gt;The Wild Ass's Skin: an electronic edition&lt;/title&gt;
&lt;editor role="translator" xml:id="EM"&gt;Ellen Marriage&lt;/editor&gt;
&lt;editor role="preface" xml:id="GS"&gt;George Saintsbury&lt;/editor&gt;
&lt;/titleStmt&gt;
&lt;editionStmt&gt;
&lt;edition n="2.0"&gt;
&lt;title&gt;Version 2.0, enriched with thematic annotations.&lt;/title&gt;
&lt;date when="2010"&gt;2010&lt;/date&gt;
&lt;/edition&gt;
&lt;respStmt&gt;
&lt;name&gt;Melissa Terras&lt;/name&gt;
&lt;resp&gt;Added thematic annotations.&lt;/resp&gt;
&lt;/respStmt&gt;
&lt;/editionStmt&gt;
&lt;extent&gt;572 Kb&lt;/extent&gt;
&lt;publicationStmt&gt;
&lt;idno type="ISBN"&gt;0-00-000000-0&lt;/idno&gt;
&lt;date when="2010-01-01"&gt;1 January 2010&lt;/date&gt;
&lt;/publicationStmt&gt;</pre>
```

```
<sourceDesc>
    <bibl><title>The Wild Ass's Skin</title> by <author>Honoré
        de Balzac</author>. <pubPlace>London</pubPlace>
        : <publisher>Dent</publisher>, <date when="1906">1906</date>. <extent>xv</extent>, <extent>28
        p.</extent>. Translated by <editor role="translator">Ellen Marriage</editor>
        ; preface by <editor role="preface">George Saintsbury</editor>.</bibl>
</sourceDesc>
</fileDesc>
<profileDesc>
    <creation>Original written in <date when="1831">1831</date>
    in <name type="city">Paris</name>.</creation>
    <langUsage>
        <language ident="en" usage="98">English</language>
        <language ident="fr" usage="1">French</language>
        <language ident="ar" usage="1">Arabic</language>
    </langUsage>
    <textClass>
        <keywords scheme="#lcsh">
            <list>
                <item>Literature</item>
                <item>Fiction and juvenile belles lettres</item>
                <item>Literature--Translations into English</item>
            </list>
        </keywords>
        <classCode scheme="#DDC">843.7</classCode>
        <catRef target="#BCS.phil" scheme="#BCS"/>
    </textClass>
    <handNotes>
        <handNote xml:id="JH" scribe="JamesHarding" script="hand" medium="ink
        .blue">handwriting in blue ink by James Harding, previous owner of the
        book</handNote>
        <handNote xml:id="ar" script="arabic">Arabic script</handNote>
    </handNotes>
</profileDesc>
<revisionDesc>
    <change when="2009-03-08" who="#MT">addition of thematic analysis</change>
    <change when="2009-03-08" who="#RvdB">addition of explanatory notes</change>
    <change when="2008-10-25" who="#RvdB">spell check</change>
```

```
<change when="2008-08-25" who="#RvdB">addition of phrase level markup</change>
<change when="2008-08-20" who="#RvdB">file creation</change>
</revisionDesc>
</teiHeader>
<text>
<!-- ... -->
</text>
</TEI>
<!-- ... -->
</teiCorpus>
```

Example 39. Encoding common metadata of a composite work in the `<teiHeader>` of a `<teiCorpus>` element.

1

SUMMARY

A complex text encoded as a `<teiCorpus>` should have a `<teiHeader>` in its own right. This TEI header on the corpus level can contain the general descriptive information about all corpus texts embedded as `<TEI>` documents. Each corpus text then should have its own `<teiHeader>`, describing only those aspects that are specific to that text.

.....

1 The TEI Guidelines provide refined ways of associating contextual information with specific (parts of) texts. See section [15.3 Associating Contextual Information with a Text](#) for more information.