

TEI By Example

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

TEI by Example. Module 2

Edward Vanhoutte

Ron Van den Branden

Melissa Terras

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

Last modified on: 2011-11-17

Licensed under a Creative Commons Attribution ShareAlike 3.0 License

TABLE OF CONTENTS

1. Introduction.....	1
2. Exploring a Minimal TEI Header.....	1
3. The TEI Header Sections.....	4
3.1 The File Description	5
3.1.1 The Title Statement	5
3.1.2 The Edition Statement.....	9
3.1.3 File Size.....	10
3.1.4 The Publication Statement	11
3.1.5 The Series Statement.....	13
3.1.6 The Notes Statement.....	14
3.1.7 The Source Description	14
3.1.8 Summary.....	18
3.2 The Encoding Description.....	21
3.2.1 The Editorial Practise Declaration.....	22
3.2.2 The Project Description.....	23
3.2.3 The Tagging Declaration.....	24
3.2.4 The Reference System Declaration.....	28
3.2.5 The Classification Declaration.....	29
3.2.6 The Metrical Notation Declaration (available in the verse module).....	31
3.2.7 The Variant Encoding (available in the textcrit module).....	33
3.2.8 Summary.....	34
3.3 The Profile Description.....	38
3.3.1 Creation.....	38
3.3.2 Language Usage.....	39
3.3.3 Text Classification.....	39
3.3.4 Document Hands.....	41
3.3.5 Summary.....	43
3.4 The Revision Description.....	44
4. The Header of a Complex Text.....	45
5. Summary.....	53
6. What's next?.....	59

Module 2: The TEI Header

1. Introduction

As will be clear by now, a document is more than its text. The TEI addresses this reality by providing formal means (elements and attributes) that allow the encoder to explicate his theory of the text in a descriptive manner. For example when a text fragment is italicised in an existing source text or should occur as such in an electronic text edited from scratch, TEI allows the encoder to express not just *that* this fragment is emphasised (by means of italics), but also *why* (because it is a title, foreign word, term, or whatever analysis the encoder wants to express).

This descriptive nature of TEI is not restricted to the actual textual contents of a document, but extends to the general meta-information one would like to associate with it. Therefore, the TEI Guidelines require that a TEI text instance be preceded by general meta-information. This ‘administrative’ meta-section is called the TEI header. While the TEI header may be intimidatingly elaborated, this tutorial module will guide you through its different sections, and point out those sections you'll most plausibly need when you start to encode texts with TEI.

1

2. Exploring a Minimal TEI Header

Let's start this section with a mental exercise (though you are free to make it as physical as you want). Before the holidays, your partner presents you with a short list of book titles she would like to read. Since it is you who took a day off early, you take this wish list and set out to the public library. Most of the titles are easy to find, except for the somewhat more cryptic entry:

Balzac or Zola (don't know exactly) ? something about a magic donkey (in English please!!!) -- sorry, dear, you're the best!

There are many ways you could approach this problem:

- flesh out all works by Zola and Balzac on the library shelves and try to find the one(s) dealing with magic donkeys
- have a look at the available titles in the 'translated literature' section

.....

1 The TEI header has a less direct relationship to the text than the actual TEI text elements. After all, the TEI header is not intended to contain actual text contents, but rather *abstractions* from the information that is related to the document, much like a library catalogue record. Moreover, the TEI header differs from most other TEI structures in that it has a more strict organisation, containing a number of mandatory elements and alternative options to encode information in a more or less formalised way. Therefore, this tutorial module will differ slightly from the others concerning the worked example, and ask a little more of your imagination.

- try to google for more information first

Depending on how greatly you value your free time, you will probably start / end up by asking the librarian, who will either scan her current knowledge of world literature or a catalogue of library records. Or, if you live in the 21st century, you will probably move to one of the library's computer terminals, search for 'Balzac' or 'Zola' in the author field, narrow the search to 'English' translations, and give it a try with 'donkey' in the title field. If you lived in the 22nd century, the search robot could probably analyse your search query, propose alternatives for unsuccessful search terms, and even suggest you'd give it a try with 'ass' instead of 'donkey'. For the time being, however, you'll have to depend on your (librarian's) world knowledge, patience, and/or creativity in order to find following information:

All Fields ▾
All formats ▾
Find

[← Back to Search results](#)
[<< Previous record](#) / [Next record >>](#)
Record 2 of 1946953

[Cite this](#)
[Email this](#)
[Add to favourites](#)

The wild ass's skin / by Honore de Balzac ; translated by Ellen Marriage ; preface by George Saintsbury

Availability

Bib ID	1944299
Format	Book
Author	Balzac, Honore de, 1799-1850
Uniform Title	Peau de chagrin. English
Description	London : Dent, 1906. xv, 288 p.
Series	Everyman's library. Fiction
Other Authors	Marriage, Ellen

Holdings	Comments	Librarian's View	Copyright Status
000	00680nam a2200229 a 4500		
001	1944299		
005	19920602000000.0		
008	910912s1906 enk 000 1 fre d		
019 1	a5881537		
035	9000005881537		
035	91125830		
040	aADFA beng cADFA dSUA		
041 1	aeng hfre		
082 0 4	a843/.7		
100 1	aBalzac, Honore de, d1799-1850.		
240 1 0	aPeau de chagrin. lEnglish		
245 1 4	aThe wild ass's skin / cby Honore de Balzac ; translated by Ellen Marriage ; preface by George Saintsbury.		
260	aLondon : bDent, c1906.		
300	axv, 288 p.		
440 0	aEveryman's library. pFiction		
700 1	aMarriage, Ellen.		
984	aANL cYY 843.7 B198		

Have a question?

For reference enquiries and general questions on items in the collection, [ask a librarian](#).

For comments and suggestions on the catalogue site, please [let us know](#).

Similar Items

- [Life of Honore de Balzac](#)
- [Honore de Balzac](#)
- [Beatrix / Honore de Balzac](#)
- [Honore de Balzac, 1799-1850](#)
- [Cesar Birotteau / Honore de Balzac](#)

Explore

- [Add to delicious](#)

Fig. 1.

It is the last field of this library catalogue that will guide you to the right library shelf and a superb holiday. This exercise vulgarises the motivation to abstract primary information about bibliographic objects into fixed categories. In the analog world, this happened on printed library catalogue records; nowadays these are entered as digital records in databases of library catalogues. These fixed categories together make up an ‘identity card’ of a literary work.

The TEI Guidelines consider such a virtual ‘identity card’ an essential part of each TEI document. It must be encoded within a **<teiHeader>** element, before the actual text contents in the **<text>** part. The ‘ID categories’ of the TEI Header are the subject of this tutorial module. As a trade-off between exhaustivity and usability, the TEI Guidelines define a wide range of specific TEI Header elements, only a few of which are mandatory. A minimal TEI header for above work would look as follows:

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>The Wild Ass's Skin: an electronic edition</title>
    </titleStmt>
    <publicationStmt>
      <p>Published as an example for the header module of TBE.</p>
    </publicationStmt>
    <sourceDesc>
      <p>Honoré de Balzac (1906). The Wild Ass's Skin.</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```

This example shows how a **<teiHeader>** element must contain a **<fileDesc>** (file description) element, providing a description of the *electronic* file. In order to be complete, it must consist of three subsections, in that order:

- **<titleStmt>**: a title statement about the electronic text
- **<publicationStmt>**: information on the publication of the electronic text
- **<sourceDesc>**: a bibliographic description of the source for the electronic text

2

The **<titleStmt>** element must minimally contain a title for the electronic text. Depending on the nature of this text, this title may repeat the original's title, followed by a paraphrase like "electronic version/transcription/edition". Details about the publication and source of the

.....

2 In a minimal **<teiHeader>**, only a description of the electronic text must be given in the **<fileDesc>** element. Such a minimal file description must consist of **<titleStmt>**, **<publicationStmt>**, and **<sourceDesc>** sections. Moreover, they must occur in this order.

electronic text in **<publicationStmnt>** and **<sourceDesc>** respectively, may consist of informal prose in loose paragraphs. More specialised elements can be used as well. These are covered in detail in the next section of this tutorial.

You will have noticed that this minimal example of a TEI header does quite a poor job providing an identity card of this novel, compared to the library record example above. However, there are two things of notice:

- the TEI header is an integral part of any TEI document, and *must* precede the **<text>** element with its actual content
- the TEI header minimally documents aspects of the title, publication, and source text of the *electronic* text

Of course, the TEI header allows for much more descriptive sophistication. The most important sections of the TEI header are treated in the next sections of this tutorial.

Summary

The TEI header contains meta-information about the electronic text, and is considered an integral part of it. Therefore, the **<teiHeader>** element must precede the **<text>** part of any TEI text, documenting at least some aspects of the electronic text in a **<fileDesc>** element. A file description minimally contains information about the title of the electronic text in **<titleStmnt>**, about its publication in **<publicationStmnt>**, and bibliographic information about the source document from which it is derived **<sourceDesc>**.

3. The TEI Header Sections

The TEI header can consist of four major parts:

- **<fileDesc>** (file description): bibliographic description of the electronic text
- **<encodingDesc>** (encoding description): description of the relation of the electronic text to its source
- **<profileDesc>** (profile description): description of the context in which the electronic text was created, and classification information
- **<revisionDesc>** (revision description): description of the revision history of the electronic text

As indicated in the previous section, the bibliographic file description (**<fileDesc>**) is the sole mandatory section of any TEI header. When other header sections are present, they must occur in above order.

3

.....

3 To ease visual recognition, the mandatory elements of TEI header (sub)sections are printed in red in the element overviews in this tutorial.

3.1 The File Description

The file description, in the **<fileDesc>** element, must occur as the first element in the TEI header. It contains a bibliographic description of the electronic text, and may consist of following subsections:

- **<titleStmt>** (title statement): groups information about the title of the electronic text and those responsible for its intellectual content
- **<editionStmt>** (edition statement): groups information relating to the edition of the electronic text
- **<extent>**: describes the approximate size of the electronic text
- **<publicationStmt>** (publication statement): groups information concerning the publication or distribution of the electronic text
- **<seriesStmt>** (series statement): groups information about the series in which an electronic text is published
- **<notesStmt>** (notes statement): collects together any notes providing information about the electronic text additional to that recorded in other parts of the bibliographic description
- **<sourceDesc>** (source description): describes the source from which an electronic text was derived or generated

Of these subsections, only the title statement (**<titleStmt>**), publication statement (**<publicationStmt>**), and source description (**<sourceDesc>**) are mandatory.

3.1.1 The Title Statement

The title statement minimally lists the title of the electronic text in a **<title>** element. Next to the title, it provides room to list detailed information about the persons or institutions responsible for different aspects of the realisation of the electronic text.

- **<title>**: contains the title for the electronic text
- **<author>**: contains the name of an/the author(s) of the electronic text
- **<editor>**: contains the name of an/the editor(s) of the electronic text
- **<sponsor>**: specifies the name of a sponsoring organisation or institution for the realisation of the electronic text
- **<funder>**: specifies the name of a party responsible for the funding of the realisation of the electronic text
- **<principal>**: supplies the name of the principal researcher responsible for the creation of an electronic text
- **<respStmt>** (statement of responsibility): supplies a statement of responsibility for the intellectual content of the electronic text, where the specialised elements for authors, editors, etc. do not suffice or do not apply

Although the electronic text can be named anything, often its title will reflect the title of the source text (if any). In order to point out the distinction, it is advised to explicate the status of the electronic text in a phrase like "an digital edition", "an electronic transcription", or the like.

The TEI Guidelines strongly advise to separate the title of an electronic text from the name of the file in which it is saved, as the latter is likely subject to change.

All elements inside **<titleStmt>** may occur as often as needed, in order to list all (sub)titles, authors, editors, or others responsible for the realisation of the electronic text. For example, the **<titleStmt>** section for our example could be expanded as follows:

```

<titleStmt>
  <title>The Wild Ass's Skin: an electronic edition</title>
  <author>Honoré de Balzac</author>
  <editor>The TBE crew</editor>
  <respStmt>
    <name>Ellen Marriage</name>
    <resp>translation</resp>
  </respStmt>
  <respStmt>
    <name>George Saintsbury</name>
    <resp>preface</resp>
  </respStmt>
  <respStmt>
    <name>Ron Van den Branden</name>
    <resp>transcription</resp>
    <resp>annotation</resp>
  </respStmt>
  <sponsor>Association for Literary and Linguistic Computing (ALLC)</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>
      <addrLine>9000 Gent</addrLine>
      <addrLine>Belgium</addrLine>
    </address>
    <email>ctb@kantl.be</email>
  </funder>
  <principal>Edward Vanhoutte</principal>
  <principal>Melissa Terras</principal>
</titleStmt>

```

4

Note the specific form of the statements of additional responsibilities inside `<respStmt>`. Each responsibility statement should contain a proper name inside `<name>`, identifying the responsible party, and describe the responsibility inside a `<resp>` element. When one person or institution has more than one responsibilities, these may be enumerated in a number of `<resp>` elements.

The example above lists the translator among the ‘additional responsibilities’ section. However, this can be understood as a kind of editor role, hence encoded as `<editor>`. In order to distinguish between different kind of editorial responsibilities, the `<editor>` element has a specific `@role` attribute, whose values can include `translator`, `editor`, `compiler`, `illustrator`, etc. Likewise, the author of the preface could be considered a kind of editor and encoded as such.

When encoding an electronic text, you might want to identify who is responsible for certain textual phenomena, such as additions, deletions, solutions of abbreviations, annotations, etc. Many of the tags for such phenomena have a `@resp` attribute, whose value should refer to an element formally identified elsewhere. Of course, the `<titleStmt>` provides an excellent location to provide such formal identification codes, by making use of the global `@xml:id` attribute. This way, the textual phenomena in a transcription can be associated directly with both the name of the responsible parties, and their roles in realising the electronic text.

The example above could thus be rephrased as follows:

.....

- 4 Note how the order of the element inside `<titleStmt>` is free. Also, the `<funder>` element illustrates how the `<titleStmt>` elements may contain common phrase level elements, for example an `<address>`.

```

<titleStmt>
  <title>The Wild Ass's Skin: an electronic edition</title>
  <author xml:id="HdB">Honoré de Balzac</author>
  <editor role="translator" xml:id="EM">Ellen Marriage</editor>
  <editor role="editor" xml:id="TBEcrew">The TBE crew</editor>
  <editor role="preface" xml:id="GS">George Saintsbury</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 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>
      <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>

```

These identifications allow the encoder to distinguish, for example, between an editorial annotation and a note by the translator in the text:

```

<!-- editorial annotation -->
<note resp="#RvdB">
  <term>ass</term>
  <gloss>donkey</gloss>
</note>
<!-- note by the translator -->
<note resp="#GS">I hesitated between "The Piece of Shagreen" and "The Wild Ass' Skin" for the title, but
Balzac's own remarks decided me. "The Magic Skin" is very weak, and "The Skin of Shagreen" hideous.</note>

```

Summary

The title statement (**<titleStmt>**) is the first mandatory subsection of the file description. It should at least contain a **<title>** element, providing a title for the electronic text. Besides the title, different parties can be identified that had been involved in the realisation of the electronic text: author (**<author>**), editor (**<editor>**), sponsor (**<sponsor>**), funder (**<funder>**), principal (**<principal>**). Other responsibilities may be encoded in a **<respStmt>** element, listing both the name of the responsible party (**<name>**), and its responsibilities in a list of **<resp>** elements. For reference purposes, it makes sense to formally identify the parties identified in the title statement with global **@xml:id** attributes.

3.1.2 The Edition Statement

The edition statement provides detailed information about the edition of the electronic text (if applicable). Similar to editions of printed texts, electronic texts may be substantially revised in different versions. Somehow closer to the world of software programs, an edition of an electronic text can be compared to the ‘release’ of a piece of software. For an electronic text, the alteration of its contents, or addition/expansion/removal of substantive (types of) meta-information could qualify a new version of an electronic text as a new edition.

The **<editionStmt>** element can contain:

- **<p>** | **<edition>**: a description of the edition; either as loose prose paragraphs (**<p>**), or in a specific **<edition>** element
- **<respStmt>**: contains descriptions of responsibilities specific to the current edition

The edition may be described either loosely in one or more paragraphs (**<p>**), or in a more specific **<edition>** element. One of both (but not both) must be present. Note that only one **<edition>** element may be used. When applicable, responsible parties and their specific responsibilities for this edition can be listed inside a **<respStmt>** element, identifying both the responsible party (**<name>**) and its responsibilities (**<resp>**).

The TEI Guidelines state that

[a]n edition statement is optional for the first release of a computer file; it is mandatory for each later release, though this requirement cannot be enforced by the parser.

[2.2.2 The Edition Statement](#)

If, for example, the digital edition of this version of *The Wild Ass's Skin* builds on an existing electronic edition, but adds a substantive new category of annotations by Melissa Terras, this could be reflected in the **<editionStmt>** as follows:

```

<editionStmt>
  <edition n="2.0">
    <title>Version 2.0, enriched with thematic annotations.</title>
    <date when="2010">2010</date>
  </edition>
  <respStmt>
    <name>Melissa Terras</name>
    <resp>Added thematic annotations.</resp>
  </respStmt>
</editionStmt>

```

Note how we can't formally identify Melissa anymore, as this person already has been identified earlier (at least, not with the same identification code, as these have to be unique within an XML document). If we want to identify her in this role, there are two solutions:

- provide a different identification code for the **@xml:id** attribute
- move the **@xml:id** identification from the **<principal>** element inside **<titleStmt>** to the **<name>** element inside the **<editionStmt>**

Summary

The particular edition of the electronic text can be described in **<editionStmt>**, either as a loose prose description in (a) paragraph(s) (**<p>**), or one **<edition>** element. Additional responsibilities associated with this edition can be stated in (a) **<respStmt>** element(s).

3.1.3 File Size

The **<extent>** section of the file description provides an analogue to the bibliographical indication of the size of printed books. It allows the encoder to express the size of the electronic text, be it in terms of its carrier medium (bits, bytes, number of diskettes/DVDs), or in terms of its contents (number of words/sentences). In this way, the TEI Guidelines aim to offer some way of formalising this often fluid notion of size in digital terms. Because of the low level of formalisation of such kinds of information, the **<extent>** element may contain a loose prose description of the amount and units of size.

For example, the size of our example text could be encoded as follows:

```
<extent>572 Kb</extent>
```

...or, in terms of its carrier medium:

```
<extent>1 5.25" floppy disk (720 Kb)</extent>
```

Summary

The **<extent>** section of the file description provides a means to record the size of the electronic text.

3.1.4 The Publication Statement

The publication statement (**<publicationStmt>**) is the second mandatory section of the file description. It provides details about the publication status of the electronic text, in one or more of following subsections:

- **<p>** | **<publisher>** | **<distributor>** | **<authority>**: description of publication details; either by means of loose prose paragraphs (**<p>**); or an identification of the publisher (**<publisher>**), the distributor (**<distributor>**), or other authority (**<authority>**) for making the electronic text available
- **<pubPlace>**: the place of publication for the electronic text
- **<address>**: the address of the publishing body of the electronic text
- **<idno>**: a standardised bibliographic identification code for the electronic text
- **<availability>**: a statement about the availability and terms of use of the electronic text
- **<date>**: the publication date of the electronic text

The publication statement can either contain a loose prose description in one or more paragraphs (**<p>**), or any of the other elements (although the TEI Guidelines recommend to use at least **<publisher>**, **<distributor>**, or **<authority>**).

The element names are quite transparent, and analogous to the labels often present in traditional bibliographic descriptions of printed works. Most of them can contain plain text and phrase-level elements, apart from **<address>**, **<availability>**, and **<idno>**. The **<address>** element must contain at least a generic **<addrLine>** element containing a single address line, or more specific address elements like **<street>**, **<name>**, **<postCode>**, or **<postBox>**. Information on availability and terms of use inside **<availability>** must be given in at least one paragraph (**<p>**). The availability of the electronic text can be typed formally in a **@status** attribute on **<availability>**, with three possible values:

- **free**: the text is freely available
- **restricted**: the text is not freely available
- **unknown**: the status of the text is unknown

The **<idno>** element can be used to provide a formal identification code in some kind of scheme. The code itself must be given in plain text; the applicable identification scheme can be identified in the value of the **@type** attribute. Possible values include ISBN ([International Standard Book Number](#)), LCCN ([Library of Congress Control Number](#)), DOI ([Digital Object Identifier](#)), etc. The **<date>** element can give the date in free prose. Additionally, for automation's sake, the use of the **@when** attribute is recommended. Its value should be a formal representation of the date, most commonly in the form yyyy-mm-dd.

5

The **<publicationStmt>** for our sample text could look as follows:

```
<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 (CTB)</name>
    <name type="institution">Royal Academy of Dutch Language and Literature</name>
    <street>Koningstraat 18</street>
    <postCode>9000</postCode>
    <name type="city">Gent</name>
    <name type="country">Belgium</name>
  </address>
  <idno type="ISBN">0-00-000000-0</idno>
  <availability status="free">
    <p>Published under a <ref target="http://creativecommons.org/licenses/by-sa/3.0/">Creative Commons
      Attribution ShareAlike 3.0 License</ref>.</p>
  </availability>
  <date when="2010-01-01">1 January 2010</date>
</publicationStmt>
```

Note how the availability description contains an external reference to an online description of the licensing scheme. This is possible since the paragraphs inside **<availability>** can contain anything a regular paragraph can contain.

Summary

The publication statement of an electronic text inside **<publicationStmt>** is the second mandatory subsection of the file description. It reflects the bibliographic publication description of a printed work, and can be provided either as loose prose in one or more paragraphs (**<p>**), or with one or more specialised elements. The TEI Guidelines advise to state at least the publisher (**<publisher>**), distributor (**<distributor>**), or any other bodies responsible for making available the electronic text (**<authority>**). Additional elements are provided for the description of the publication place (**<pubPlace>**), publication address (**<address>**), bibliographic identification code (**<idno>**), availability and terms of use (**<availability>**), and publication date (**<date>**).

.....

- 5 For a complete list of allowed date expressions in the **<date>** element's **@when** attribute: see [W3C XML Schema Part 2: Datatypes Second Edition](#).

3.1.5 The Series Statement

If the electronic text is published in a series, this series can be described in the `<seriesStmt>` element. It may contain following elements:

- `<p>` | `<title>`: a description of the series, either in a loose prose description (`<p>`), or by naming the series inside `<title>`
- `<idno>`: a standardised bibliographic identification code for the series in which the electronic text is published
- `<respStmt>`: statement of responsibility for the realisation of the series

The series statement may either be given in loose prose inside paragraphs (`<p>`), or must at least name the title of the series (`<title>`). Additionally, an identification code for the series, and/or for the electronic text within the series, can be given inside `<idno>` with an appropriate value for its `@type` attribute. Responsible parties for the realisation of the series can be listed in (a) `<respStmt>` element(s).

For example, our sample electronic text could be published in a series that could be described as follows:

```
<seriesStmt>
  <title>The TBE collection: sample texts encoded with TEI.</title>
  <idno type="ISSN">0000-0001</idno>
  <idno type="installment">1</idno>
  <respStmt>
    <name>Edward Vanhoutte</name>
    <resp>compiler</resp>
  </respStmt>
</seriesStmt>
```

Note how the second `<idno>` element is used to identify the electronic text within the series. The `@type` attribute indicates here that the identification refers to the sequence number of the instalments in the series. It could, of course, indicate other reference schemes as well (such as volumes, issues, ...).

Summary

Details on the series in which an electronic text was published may be recorded in the `<seriesStmt>` element. The series statement may either be given as loose prose inside paragraphs (`<p>`), or must at least name the title of the series (`<title>`). Additionally, an identification code for the series, and/or for the electronic text within the series, can be given inside `<idno>` with an appropriate value for its `@type` attribute. Responsible parties for the realisation of the series can be listed in (a) `<respStmt>` element(s).

3.1.6 The Notes Statement

The `<notesStmt>` section of the file description is reserved for additional information that is not covered in the general bibliographic description. Each piece of additional information should be encoded in a separate `<note>`:

```
<notesStmt>
  <note>OCR scanning done at KANTL, Gent.</note>
</notesStmt>
```

Summary

The `<notesStmt>` section of the file description is reserved for additional information that is not covered in the general bibliographic description. Each piece of additional information should be encoded in a separate `<note>`

3.1.7 The Source Description

The source description inside `<sourceDesc>` is the third required subsection of the file description. It should contain one of following elements:

- `<p>` | `<bibl>` | `<biblStruct>` | `<biblFull>` | `<listBibl>`: the bibliographic description of the source text, either as a loose prose description in a paragraph (`<p>`); a formal bibliographic description, either loose (`<bibl>`), structured (`<biblStruct>`), or exhaustive (`<biblFull>`); or a list of bibliographic references (`<listBibl>`)

The source text of the electronic text can be described either as loose prose in (a) paragraph(s) (`<p>`), or by means of a more specialised bibliographical element (`<bibl>`, `<biblStruct>`, `<biblFull>`, `<listBibl>`).

Of course, not all texts are derived from a material source text. In fact, lots of TEI documents are encoded from scratch, just like regular text files produced with text processing software. For such texts, a kind of ‘dummy’ statement can be given in a paragraph inside `<sourceDesc>`. For example, the `<sourceDesc>` element of this TBE tutorial module (a native TEI electronic document) looks as follows:

```
<sourceDesc>
  <p>No source, born digital.</p>
</sourceDesc>
```

If possible, however, it is recommended to bibliographically describe the material source document using a more specialised TEI element for bibliographic description. The `<bibl>`,

<biblStruct>, and **<biblFull>** elements share a common set of allowed child elements, but differ in their degree of completeness and strictness. The most informal of the specialised bibliographical elements is **<bibl>**, which allows a prose-like bibliographic description, possibly interspersed with bibliographic elements, the most important of which are:

- **responsibilities:**
 - **<author>**: the author of the source text
 - **<editor>**: the editor of the source text
 - **<distributor>**: the distributing agency of the source text
 - **<publisher>**: the publishing agency of the source text
 - **<funder>**: the funding agency of the source text
 - **<principal>**: the principal researcher responsible for the realisation of the source text
 - **<sponsor>**: the sponsoring agency of the source text
 - **<respStmnt>**: other responsibilities for the source text
- **edition:**
 - **<title>**: the title of the source text
 - **<date>**: the publication date of the source text
 - **<pubPlace>**: the publication place of the source text
 - **<edition>**: the edition of the source text
 - **<series>**: the series in which the source text was published
 - **<idno>**: a bibliographic reference code for the source text
 - **<biblScope>**: the scope of the bibliographic reference of the source text
 - **<extent>**: the size of the source text

The **<sourceDesc>** for our example could look as follows:

```
<sourceDesc>
  <p>The Wild Ass's Skin by Honoré de Balzac. London : Dent, 1906. xv, 288 p. Translated by Ellen Marriage ;
  preface by George Saintsbury.</p>
</sourceDesc>
```

... or with a **<bibl>** element:

```
<sourceDesc>
  <bibl>The Wild Ass's Skin by Honoré de Balzac. London : Dent, 1906. xv, 288 p. Translated by Ellen
  Marriage ; preface by George Saintsbury.</bibl>
</sourceDesc>
```

... more formally:

```

<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>288 p.</extent> Translated by <editor role="translator">Ellen Marriage</editor>
  ; preface by <editor role="preface">George Saintsbury</editor>.</bibl>
</sourceDesc>

```

The same information can be structured more rigorously using the **<biblStruct>** element. A structured bibliography may contain the same bibliographic elements, but structured in a more explicit way on three possible levels:

- **<analytic>**: bibliographic description of an item published within a monograph or journal:
 - **<title>**: the title of the article or contribution
 - **<author>**: the author of the article or contribution
 - **<editor>**: the editor of the article or contribution
 - **<respStmt>**: other responsibilities for the article or contribution
- **<monogr>**: bibliographic description of an item published as an independent item:
 - **<author>**: the author of the monograph
 - **<editor>**: the editor of the monograph
 - **<respStmt>**: other responsibilities for the monograph
 - **<title>**: the title of the monograph
 - **<imprint>**: information regarding the publication of the monograph: **<biblScope>**, **<distributor>**, **<pubPlace>**, **<publisher>**, **<date>**
 - **<biblScope>**: the bibliographic scope of the monograph (volume, issue)
 - **<note>**: additional bibliographical notes for the monograph
- **<series>**: bibliographic description of the series in which a work has been published:
 - **<title>**: the series' title
 - **<editor>**: the series' editor
 - **<respStmt>**: other responsibilities for the series
 - **<biblScope>**: the bibliographic scope for the bibliographic item within the series

Our example could be elaborated as follows:

```

<sourceDesc>
  <biblStruct>
    <monogr>
      <title>The Wild Ass's Skin</title>
      <author>Honoré de Balzac</author>
      <editor role="translator">Ellen Marriage</editor>
      <editor role="preface">George Saintsbury</editor>
      <imprint>
        <pubPlace>London</pubPlace>
        <publisher>Dent</publisher>
        <date when="1906">1906</date>
      </imprint>
      <extent>xv</extent>
      <extent>288 p.</extent>
    </monogr>
  </biblStruct>
</sourceDesc>

```

The **<biblFull>** element requires the most extensive bibliographic description for the source of the electronic text, organised in the same categories as the file description of the electronic text itself (without the **<sourceDesc>** section, of course): a mandatory title statement (**<fileDesc>**), optional edition statement (**<editionStmt>**), indication of the size (**<extent>**), mandatory publication statement (**<publicationStmt>**), series statement (**<seriesStmt>**), and possibly additional bibliographic notes (**<notesStmt>**). As this level of detail exceeds the aims of this introductory tutorial, you are kindly referred to the **<biblFull>** reference section of the TEI Guidelines for a full reference and examples.

If an electronic text is derived from more source texts, these can all be described with the desired granularity using **<bibl>**, **<biblStruct>**, and **<biblFull>**. When doing so, the **<listBibl>** provides a convenient way to group these bibliographic descriptions.

Although our example text is derived from only one source, following example illustrates how **<listBibl>** can be used:

```

<sourceDesc>
  <listBibl>
    <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>288 p.</extent> Translated by <editor role="translator">Ellen
    Marriage</editor> ; preface by <editor role="preface">George Saintsbury</editor>.</bibl>
    <!-- description of other sources goes here -->
    <bibl>
      <!-- .. -->
    </bibl>
  </listBibl>
</sourceDesc>

```

6

Summary

The description of the material sources for an electronic text inside **<sourceDesc>** is the third mandatory subsection of the file description. It must contain either a prose description of the source in (a) paragraph(s) (**<p>**), or a more formalised description using one of the specific TEI elements for bibliographic description: **<bibl>**, **<biblStruct>**, or **<biblFull>**. When an electronic text is derived from more source texts, these descriptions may be grouped inside a **<listBibl>** element.

3.1.8 Summary

When all pieces are put together, the file description for our example can look as follows:

.....

- 6 When the dedicated **msdescription** TEI module for Manuscript Description is included in a TEI schema, the **<sourceDesc>** section of the TEI header provides a specific element for the bibliographic description of the source manuscript for the electronic text: **<msDesc>**. Due to the extensiveness of this element and the specificity of its use, you are referred to the TEI Guidelines for a full reference on how to bibliographically describe source manuscripts. Chapter [10 Manuscript Description](#) is nearly completely devoted to this element.

```

<fileDesc>
  <titleStmt>
    <title>The Wild Ass's Skin: an electronic edition</title>
    <author xml:id="HdB">Honoré de Balzac</author>
    <editor role="translator" xml:id="EM">Ellen Marriage</editor>
    <editor role="editor" xml:id="TBEcrew">The TBE crew</editor>
    <editor role="preface" xml:id="GS">George Saintsbury</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 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>
        <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>
  <editionStmt>
    <edition n="2.0">
      <title>Version 2.0, enriched with thematic annotations.</title>
      <date when="2010">2010</date>
    </edition>
    <respStmt>
      <name>Melissa Terras</name>
      <resp>Added thematic annotations.</resp>
    </respStmt>
  </editionStmt>
  <extent>572 Kb</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>
  </publicationStmt>

```

```

<pubPlace>Gent</pubPlace>
<address>
  <name type="institution">Centre for Scholarly Editing and Document Studies (CTB)</name>
  <name type="institution">Royal Academy of Dutch Language and Literature</name>
  <street>Koningstraat 18</street>
  <postCode>9000</postCode>
  <name type="city">Gent</name>
  <name type="country">Belgium</name>
</address>
<idno type="ISBN">0-00-000000-0</idno>
<availability status="free">
  <p>Published under a <ref target="http://creativecommons.org/licenses/by-sa/3.0/">Creative Commons
  Attribution ShareAlike 3.0 License</ref>.</p>
</availability>
<date when="2010-01-01">1 January 2010</date>
</publicationStmt>
<seriesStmt>
  <title>The TBE collection: sample texts encoded with TEI.</title>
  <idno type="ISSN">0000-0001</idno>
  <idno type="installment">1</idno>
  <respStmt>
    <name>Edward Vanhoutte</name>
    <resp>compiler</resp>
  </respStmt>
</seriesStmt>
<notesStmt>
  <note>OCR scanning done at KANTL, Gent.</note>
</notesStmt>
<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>288 p.</extent>. Translated by <editor role="translator">Ellen
  Marriage</editor> ; preface by <editor role="preface">George Saintsbury</editor>.</bibl>
</sourceDesc>
</fileDesc>

```

Summary

The file description (**<fileDesc>**) is the first and sole mandatory section of the TEI header. It contains a description of the electronic text, in a mandatory title statement (**<titleStmt>**), a description of the specific edition of the electronic text is published (**<editionStmt>**), the file size (**<extent>**), a mandatory

description of publication details of the electronic text (**<publicationStmt>**), a description of the series in which the electronic text is published (**<seriesStmt>**), additional bibliographic notes (**<notesStmt>**), and a mandatory bibliographic description of the electronic text's material source text (**<sourceDesc>**).

3.2 The Encoding Description

The encoding description, in the **<encodingDesc>** element, is the second major section of the TEI header. It documents the relationship between the electronic text and its source text, either as loose prose in (a) paragraph(s), or in minimally one of more specific elements. Some of these specific elements provide details on the editorial principles for the transcription, and/or the project in which the electronic text originated:

- **<editorialDecl>** (editorial practise declaration): description of aspects of the editorial practise that informed the creation of the electronic text
- **<projectDesc>** (project description): description of the aims and circumstances of project that informed the creation of the electronic text

Besides these descriptive subsections, the encoding declaration is the place where reference systems are defined or declared that can be used anywhere in the document:

- **<tagsDecl>** (tagging declaration): information about the tags used for the encoding of the electronic text
- **<refsDecl>** (reference system declaration): declaration of reference systems used in the encoding of the electronic text
- **<classDecl>** (classification declaration): declaration of classification scheme(s) used to classify the electronic text elsewhere in the document

Finally, the encoding description can contain subsections that are only enabled when specific TEI modules are included in the TEI schema:

- **poetry** (see [TBE Module 4: Poetry](#)):
 - **<metDecl>** (metrical notation declaration): declaration of the notation for metrical analyses of poetry
- **critical apparatus** (see [TBE Module 7: Scholarly Editing](#))
 - **<variantEncoding/>**: declares the method used to encode text-critical variants

7

.....

7 Besides these elements, the encoding description can contain other subsections as well, dealing with, for example, a sampling declaration for text collections (**<samplingDecl>**), a declaration of nonstandard characters (**<charDecl>**), the applications used for processing the electronic text (**<appInfo>**), etc. Due to their specificity, these elements are not discussed in this introductory tutorial; see section [2.3 The Encoding Description](#) of the TEI Guidelines for full coverage.

3.2.1 The Editorial Practise Declaration

The editorial policy used when marking up the electronic text can be documented in `<editorialDecl>`, either as loose prose in (a) paragraph(s), or in minimally one specific element. Following elements are most common:

- **<correction>**: describes if / how / when corrections have been made in the text. A **@status** attribute can indicate the degree of correction applied to the text (low, medium, high, or unknown); a **@method** attribute can formalise whether corrections have been applied `silently` or by means of markup.
- **<normalization>**: describes if / how / when the text has been normalised. A **@source** attribute can point to the description of the authority for the normalisations; a **@method** attribute can formalise whether normalisations have been applied `silently` or by means of markup.
- **<quotation>**: describes how quotation marks in the original have been treated in the electronic text. A **@marks** attribute can record the degree to which quotation marks have been retained in the electronic text (`none`, `some`, or `all`).
- **<hyphenation>**: describes how hyphenated text in the original has been treated in the electronic text. An **@eol** attribute can record the degree to which end-of-line hyphenation has been retained in the electronic text (`none`, `some`, `all`, or `hard` (only hard end-of-line hyphenation has been retained)).
- **<interpretation>**: describes what interpretive information has been added to the text, apart from the transcription

All of these elements must contain at least one paragraph (`<p>`) containing the description.

For our example, the `<editorialDecl>` subsection could look as follows:

```

<editorialDecl>
  <correction method="markup">
    <p>Apparent errors have been corrected using the <sic> / <corr> elements, wrapped in a <choice>
    element.</p>
  </correction>
  <normalization method="markup" source="http://www.oed.com/">
    <p>Spelling has been modernised using the <orig> / <reg> elements, wrapped in a <choice> 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>

```

Summary

The **<editorialDecl>** subsection of the encoding description documents the editorial practise that has been adopted for the encoding of the electronic text. It may consist of either a loose prose in (a) paragraph(s), or more specialised elements describing the editorial policy concerning corrections (**<correction>**), normalisation (**<normalization>**), quotation (**<quotation>**), hyphenation (**<hyphenation>**), and interpretation (**<interpretation>**).

3.2.2 The Project Description

The aims and purposes for which the electronic text has been created can be given in **<projectDesc>**, as well as any other information regarding this endeavour. The structure of **<projectDesc>** is simple: it consists simply of (a) paragraph(s) (**<p>**).

For example:

```

<projectDesc>
  <p>Text encoded for <soCalled>The TBE collection: sample texts encoded with TEI</soCalled>, aiming at
  providing a collection of prime exemplar TEI encoded materials.</p>
</projectDesc>

```

Summary

The **<projectDesc>** subsection of the encoding description provides more information about the aims and goals for which the electronic text has been created. This is provided as a prose description in one or more paragraphs (**<p>**).

3.2.3 The Tagging Declaration

The XML elements that have been used to mark up the text can be formally documented in the **<tagsDecl>** subsection of the encoding description. Following aspects can be documented:

- **<namespace>**: formally identifies the namespace to which the XML elements belong that are documented in its **<tagUsage>** children
- **<rendition>**: declares a rendering style for one or more XML elements in the electronic text

The XML tags occurring within an electronic text can be documented with **<tagUsage>** elements, grouped within a **<namespace>** element. The **<namespace>** element must include a formal reference to the namespace of these XML elements in its **@name** attribute. For any default TEI text, this namespace reference should point to the `http://www.tei-c.org/ns/1.0` namespace definition. Each distinct XML element occurring within the **<text>** part of the electronic text should be documented in its own **<tagUsage>** element, providing a prose description for the use of this element in the electronic text. The element name must be provided in the **@gi** (general identifier) attribute. Additionally, the number of occurrences can be recorded in an **@occurs** attribute, and the number of occurrences with a unique identification code can be given in a **@withId** attribute. A **@render** attribute can point to a standard rendering style for this element:

```
<tagsDecl>
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="div" occurs="300" withId="300" render="#division">Marks text divisions in the source
    text.</tagUsage>
    <tagUsage gi="p" occurs="8302" withId="8300" render="#paragraph">Marks paragraphs in the source
    text.</tagUsage>
    <!-- <tagUsage> elements for all other distinct elements -->
  </namespace>
</tagsDecl>
```

8

.....

8 Note how the listing of the distinct elements in an electronic text and their occurrences (with ID code) can only be provided after the completion of the encoding, before publication. The counting of all unique text

Of course, if your electronic document contains elements from other namespaces, these should be documented within their dedicated `<namespace>` element. Note how the `@render` attribute points to the definition of a rendering style somewhere else in the document. The tagging declaration is the place for such definitions as well, by means of different `<rendition>` elements for each distinct rendition style. Rendition styles can be hand-crafted, or refer to existing styling languages such as [CSS](#) (Cascading Style Sheets) or [XSL FO](#) (eXtensible Stylesheet Language: Formatting Objects). The `@scheme` attribute defines one of these styles: `css` (CSS), `xslfo` (XSL FO), `free` (informal free text description), or `other` (any other formal rendition scheme). The contents of the `<rendition>` element then can provide the formal rendition rules expressed in any of the schemes identified. For example, the styles ‘div’ and ‘p’ can be defined in terms of CSS rules as follows:

```
<tagsDecl>
  <rendition scheme="css" xml:id="division">display:block; margin: 1em;</rendition>
  <rendition scheme="css" xml:id="paragraph">display:block; margin-bottom: 0.5em;</rendition>
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="div" occurs="300" withId="300" render="#division">Marks text divisions in the source
    text.</tagUsage>
    <tagUsage gi="p" occurs="8302" withId="8300" render="#paragraph">Marks paragraphs in the source
    text.</tagUsage>
  </namespace>
</tagsDecl>
```

Note how the `@xml:id` value of the `<rendition>` elements is used to refer to these definitions with the `@render` attribute. Furthermore, these rendering styles can be referred to with any TEI element in the electronic text by means of the global `@rendition` attribute. For example, the main title on the title page:

.....

elements and their occurrences is typically a task that can be automated, for example by using [XSLT](#) or [XSLT 2.0](#) stylesheets. The TEI Wiki has a dedicated section with useful XSLT snippets: if you can't figure out how to perform a specific XSLT job for your TEI files, the [XSLT section](#) on the TEI Wiki may be a good place to start looking for inspiration.

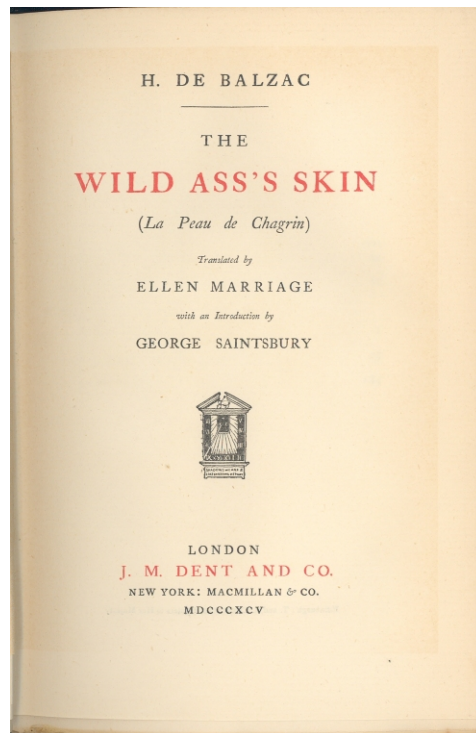


Fig. 2.

...appearing as large red text in small caps, can be encoded as follows:

```

<TEI>
  <teiHeader>
    <!-- ... -->
    <encodingDesc>
      <!-- ... -->
      <tagsDecl>
        <rendition scheme="css" xml:id="division">display:block; margin: 1em;</rendition>
        <rendition scheme="css" xml:id="paragraph">display:block; margin-bottom: 0.5em;</rendition>
        <rendition scheme="css" xml:id="red">color:red;</rendition>
        <rendition scheme="css" xml:id="smallcaps">font-variant:small-caps;</rendition>
        <rendition scheme="css" xml:id="large">font-size:large;</rendition>
        <!-- ... -->
        <namespace name="http://www.tei-c.org/ns/1.0">
          <tagUsage gi="div" occurs="300" withId="300" render="#division">Marks text divisions in the
            source text.</tagUsage>
          <tagUsage gi="p" occurs="8302" withId="8300" render="#paragraph">Marks paragraphs in the
            source text.</tagUsage>
          <!-- ... -->
        </namespace>
      </tagsDecl>
      <!-- ... -->
    </encodingDesc>
    <!-- ... -->
  </teiHeader>
  <text>
    <front>
      <titlePage>
        <!-- ... -->
        <title type="main">The <seg rendition="#red #smallcaps #large">Wild Ass's Skin</seg></title>
        <!-- ... -->
      </titlePage>
    </front>
    <!-- ... -->
  </text>
</TEI>

```

Summary

The **<tagsDecl>** subsection of the encoding declaration can document all tags, their usage and rendition in the electronic document. Specific rendition styles can be defined with a **<rendition>** element, whose **@scheme** attribute identifies the formal rendition scheme. Documentation for all

unique tags occurring inside the electronic document's **<text>** element should be grouped per namespace to which they belong, within a **<namespace>** element. Its **@name** attribute must point to a formal definition of that namespace. Each unique tag of that namespace then can be documented with a dedicated **<tagUsage>** element, containing a prose description, the tag's name in the **@gi** attribute, and indications for its occurrence within the electronic document, either in general (the **@occurs** attribute), or with a unique identification code (**@withId**).

3.2.4 The Reference System Declaration

Any reference schemes that are used in the electronic text can be declared in the **<refsDecl>** subsection of the encoding description. They can be defined either as loose prose descriptions in (a) paragraph(s) (**<p>**), or with specialised elements. Because of the complexity of these specialised elements, this tutorial section only treats the informal prose description.

9

The reference system declaration may be used to document any reference system used in the electronic text, for example the numbering schemes in **@n** attributes of certain elements, the composition of **@xml:id** values for certain elements, and so on.

For example, the numbering scheme of paragraphs, and identification codes of chapters could be documented as follows in the example text:

```
<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>
```

Summary

Any reference scheme used in the electronic text can be documented in the **<refsDecl>** subsection of the encoding declaration. The description can happen either informally in (a) paragraph(s), or more formally in dedicated TEI elements (not treated in this introductory tutorial).

.....

9 The reference schemes used in the document can be declared very formally with specific elements within **<refsDecl>**. Because this is much too complex for this introductory tutorial, you are referred to section [2.3.5 The Reference System Declaration](#) of the TEI Guidelines for full coverage.

3.2.5 The Classification Declaration

If you want to classify the electronic text using some kind of classification scheme or taxonomy, this taxonomy should be defined inside the `<classDecl>` subsection of the encoding description. The actual classification of the text is done in another part of the TEI header (see [3.3.3 Text Classification \(p. 39\)](#)), but it must point to one of the taxonomies defined here. The classification schemes used in the electronic document must each be defined in a dedicated element:

- `<taxonomy>`: defines a typology used to classify texts

The `<taxonomy>` element can either refer to an existing classification scheme, or define an own classification scheme, and should be formally identified in an `@xml:id` attribute. If the taxonomy refers to an existing classification scheme, this should be described in a `<bibl>` element. The library record for our example text contains a reference to the [Dewey Decimal Classification](#) (DDC) scheme (see code 082 in the screenshot above). If we want to include this classification code, and the [Library of Congress Subject Headings](#) in our electronic version of the text, these schemes should be referred to in a `<taxonomy>` element as follows:

```
<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="lcsch">
    <bibl>
      <title>Library of Congress Subject Headings</title>
    </bibl>
  </taxonomy>
</classDecl>
```

If the classification scheme is less universal, or if you want to roll your own, the `<taxonomy>` element can be used as well. Apart from an optional bibliographical reference in `<bibl>`, the classification categories can be defined in separate `<category>` elements, each with their own `@xml:id` identification code. The category can be described in a `<catDesc>` element. As the classification categories can nest, it is possible to define hierarchical classification systems. For example, it could make sense to classify this novel in the terms of the Balzac's own [plan of the *Comédie Humaine*](#), which the author envisaged as the encompassing series for his complete prose oeuvre:

```

<taxonomy xml:id="BCS">
  <category xml:id="BCS.man">
    <catDesc>Studies of Manners</catDesc>
    <category xml:id="BCS.man.priv">
      <catDesc>Scenes from Private Life</catDesc>
    </category>
    <category xml:id="BCS.man.prov">
      <catDesc>Scenes from provincial life</catDesc>
      <category xml:id="BCS.man.prov.cel">
        <catDesc>The Celibates</catDesc>
      </category>
      <category xml:id="BCS.man.prov.par">
        <catDesc>Parisians in the Country</catDesc>
      </category>
      <category xml:id="BCS.man.prov.jeal">
        <catDesc>The Jealousies of a Country Town</catDesc>
      </category>
    </category>
    <category xml:id="BCS.man.par">
      <catDesc>Scenes from Parisian life</catDesc>
      <category xml:id="BCS.man.thir">
        <catDesc>The Thirteen</catDesc>
      </category>
      <category xml:id="BCS.man.rel">
        <catDesc>Poor Relations</catDesc>
      </category>
    </category>
    <category xml:id="BCS.man.pol">
      <catDesc>Scenes from political life</catDesc>
    </category>
    <category xml:id="BCS.man.mil">
      <catDesc>Scenes from military life</catDesc>
    </category>
    <category xml:id="BCS.man.cou">
      <catDesc>Scenes from country life</catDesc>
    </category>
  </category>
  <category xml:id="BCS.phil">
    <catDesc>Philosophical studies</catDesc>
  </category>
  <category xml:id="BCS.ana">
    <catDesc>Analytical studies</catDesc>
  </category>
</taxonomy>

```

```
</category>
</taxonomy>
```

In this example, a separate taxonomy is created for Balzac's *Comédie Humaine*, indicated with the identification code BCS. It consists of 6 subcategories, each in its own **<category>** element, and a more detailed value for its **@xml:id** attribute. Some of these categories contain even further categories. These categories can be referred to in the actual text classification further in the TEI header (see [3.3.3 Text Classification \(p. 39\)](#)).

Summary

If the TEI header contains a formal text classification, the classification schemes used must be defined in the **<classDecl>** subsection of the encoding description. Each classification scheme should be identified by means of the **@xml:id** attribute of a **<taxonomy>** element. Such taxonomy declarations can either refer to public classification schemes, with a **<bibl>** element, or define its own classification categories inside specific **<category>** elements. Such category descriptions should describe the category in a **<catDesc>** element.

3.2.6 The Metrical Notation Declaration (available in the verse module)

When the TEI module **verse** is included in the TEI scheme, the encoding description contains an additional element for the declaration of the metrical notation used in the analysis of poetry: **<metDecl>**.

Crosslink

See

- TBE module [4: Poetry](#) for a discussion of the elements in the TEI **verse** module
- TBE module [8: Customising TEI, ODD, Roma](#) for a discussion of how to include the TEI **verse** module in a TEI schema

The metrical notation may be defined either informally, in a (series of) paragraph(s) (**<p>**), or formally using one or more **<metSym>** elements. An informal declaration may look as follows:

```
<metDecl>
  <p>The classical scansion system has been used, which marks quantitative metre originally by a macron (here a dash '-') for long syllables and a breve (here a 'u') for short syllables. A bar '|' is used to mark the foot boundary and a slash '/' marks the line boundary.</p>
</metDecl>
```

This system can be declared more formally via one or more **<metSym>** (metrical notation symbol) elements. Each metrical symbol to be used in an analysis in the electronic text must be defined in the **@value** attribute, and described in the text contents of a **<metSym>** element. The previous example could be formalised as follows:

```
<metDecl>
  <metSym value="-">long syllable</metSym>
  <metSym value="u">short syllable</metSym>
  <metSym value="|">foot boundary</metSym>
  <metSym value="/">line boundary</metSym>
</metDecl>
```

After having declared the notation system for metrical analysis, you can use this notation system for metrical analyses in the electronic text. For example, if you consider following section in the text of *The Wild Ass's Skin* a poem:

```
<lg type="poem" rendition="#center #caps">
  <l>Possessing me thou shalt possess all things,</l>
  <l>but thy life is mine, for God has so willed it.</l>
  <l>Wish, and thy wishes shall be fulfilled;</l>
  <l>but measure thy desires, according</l>
  <l>to the life that is in thee.</l>
  <l>This is thy life,</l>
  <l>with each wish I must shrink</l>
  <l>even as thy own days.</l>
  <l>Wilt thou have me? Take me.</l>
  <l>God will hearken unto thee.</l>
  <l>So be it!</l>
</lg>
```

You can enrich the transcription with a metrical analysis by means of the specific **@met** attribute. It should contain the symbols for the metrical notation system you declared in the **<metDecl>** subsection of the encoding description:

```

<lg type="poem" rendition="#center #caps">
  <l met="uuu-|uuuuuu/">Possessing me thou shalt possess all things,</l>
  <l met="uu-u-|uuuuuu/">but thy life is mine, for God has so willed it.</l>
  <l met="u|u-uuu-uu/">Wish, and thy wishes shall be fulfilled;</l>
  <l met="u-uuu-|uuu/">but measure thy desires, according</l>
  <l met="uu-|uuu-/">to the life that is in thee.</l>
  <l met="uu-/">This is thy life,</l>
  <l met="u-u-uu/">with each wish I must shrink</l>
  <l met="-uu-u-/">even as thy own days.</l>
  <l met="u-u-|-u/">Wilt thou have me? Take me.</l>
  <l met="uuuuuu-/">God will hearken unto thee.</l>
  <l met="--u/">So be it!</l>
</lg>

```

Summary

When the **verse** TEI module is included in a TEI scheme, the encoding description contains a specific subsection for the declaration of metrical notation systems: **<metDecl>**. It can contain either an informal prose description of such a system in (a) paragraph(s) (**<p>**), or make use of more formalised **<metSym>** (metrical notation symbol) elements. Their **@value** attributes must specify a symbol for a metrical phenomenon that is described as their text contents. This metrical notation system can then be used in the specific **@met** attribute on poetic structures in the electronic text.

3.2.7 The Variant Encoding (available in the **textcrit** module)

When the TEI module **textcrit** is included in the TEI scheme, the encoding description contains an additional element for the declaration of the method used to indicate text-critical variants: **<variantEncoding/>**.

Crosslink

See

- TBE module [7: Scholarly Editing](#) for a discussion of the elements in the TEI **textcrit** module
- TBE module [8: Customising TEI, ODD, Roma](#) for a discussion of how to include the TEI **textcrit** module in a TEI schema

The **<variantEncoding/>** element is an empty element with two mandatory attributes. With the **@method** attribute, you must identify one of three methods for the encoding of text-critical variants in the electronic text:

- **location-referenced**: apparatus entries are anchored to identified locations in the text

- `double-end-point`: apparatus entries are anchored to the precise start and end point of the lemma in a base text
- `parallel-segmentation`: apparatus entries contain all text variants as alternative readings

For a full reference of these systems, see chapter [12 Critical Apparatus](#) of the TEI Guidelines.

The second aspect of the system for the encoding of text-critical variants that must be indicated, is the location of the text-critical apparatus. This must be done in the **@location** attribute, with two possible values:

- `internal`: the text-critical apparatus is encoded within the running text
- `external`: the text-critical apparatus is encoded outside the running text

If we wanted to create a digital text-critical edition of *The Wild Ass's Skin* by collating different editions of the novel, we should include the **textcrit** module in our TEI schema and declare the system used to represent the textual variation in `<variantEncoding/>`. The following declaration, for example, specifies that the textual variation is encoded in the running text, using the `parallel-segmentation` method:

```
<variantEncoding method="parallel-segmentation" location="internal"/>
```

For a full treatment of recording textual variation in critical editions, see TBE module [7. Scholarly Editing](#).

Summary

When the **textcrit** TEI module is included in a TEI scheme, the encoding description contains a specific subsection for the declaration of the encoding system for textual variation: `<variantEncoding/>`. It is an empty element that must have two attributes. The **@method** attribute indicates which of three methods is used to encode textual variation (`location-referenced`, `double-end-point`, or `parallel-segmentation`). The **@location** attribute specifies where the critical apparatus is located: `internal` or `external` to the base text.

3.2.8 Summary

In the previous sections, we added the encoding description for our electronic edition of *The Wild Ass's Skin*, describing the editorial principles; the aim and purposes of the encoding; the different XML elements, their use and rendition; the system(s) that will be used to classify the text (further in the header), the notation system for metrical analyses of poems in the text, and the method of recording textual variation for the text-critical edition. This amounts to following encoding description:

```

<encodingDesc>
  <editorialDecl>
    <correction method="markup">
      <p>Apparent errors have been corrected using the <sic> / <corr> elements, wrapped in a <choice>
        element.</p>
    </correction>
    <normalization method="markup" source="http://www.oed.com/">
      <p>Spelling has been modernised using the <orig> / <reg> elements, wrapped in a <choice> 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 TEI</soCalled>, aiming at
      providing a collection of prime exemplar TEI encoded materials.</p>
  </projectDesc>
  <tagsDecl>
    <rendition scheme="css" xml:id="division">display:block; margin: 1em;</rendition>
    <rendition scheme="css" xml:id="paragraph">display:block; margin-bottom: 0.5em;</rendition>
    <rendition scheme="css" xml:id="red">color:red;</rendition>
    <rendition scheme="css" xml:id="smallcaps">font-variant:small-caps;</rendition>
    <rendition scheme="css" xml:id="large">font-size:large;</rendition>
    <!-- ... -->
    <namespace name="http://www.tei-c.org/ns/1.0">
      <tagUsage gi="div" occurs="300" withId="300" render="#division">Marks text divisions in the source
        text.</tagUsage>
      <tagUsage gi="p" occurs="8302" withId="8300" render="#paragraph">Marks paragraphs in the source
        text.</tagUsage>
      <!-- ... -->
    </namespace>
  </tagsDecl>
  <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="lsh">

<bibl>

<title>Library of Congress Subject Headings</title>

</bibl>

</taxonomy>

<taxonomy xml:id="BCS">

<category xml:id="BCS.man">

<catDesc>Studies of Manners</catDesc>

<category xml:id="BCS.man.priv">

<catDesc>Scenes from Private Life</catDesc>

</category>

<category xml:id="BCS.man.prov">

<catDesc>Scenes from provincial life</catDesc>

<category xml:id="BCS.man.prov.cel">

<catDesc>The Celibates</catDesc>

</category>

<category xml:id="BCS.man.prov.par">

<catDesc>Parisians in the Country</catDesc>

</category>

<category xml:id="BCS.man.prov.jeal">

<catDesc>The Jealousies of a Country Town</catDesc>

</category>

</category>

<category xml:id="BCS.man.par">

<catDesc>Scenes from Parisian life</catDesc>

<category xml:id="BCS.man.thir">


```

        <catDesc>The Thirteen</catDesc>
    </category>
    <category xml:id="BCS.man.rel">
        <catDesc>Poor Relations</catDesc>
    </category>
</category>
    <category xml:id="BCS.man.pol">
        <catDesc>Scenes from political life</catDesc>
    </category>
    <category xml:id="BCS.man.mil">
        <catDesc>Scenes from military life</catDesc>
    </category>
    <category xml:id="BCS.man.cou">
        <catDesc>Scenes from country life</catDesc>
    </category>
</category>
    <category xml:id="BCS.phil">
        <catDesc>Philosophical studies</catDesc>
    </category>
    <category xml:id="BCS.ana">
        <catDesc>Analytical studies</catDesc>
    </category>
</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>

```

Summary

The encoding description (**<encodingDesc>**) is the second section of the TEI header. It describes the relationship between the electronic text and its source text, either as loose prose in (a) paragraph(s), or in minimally one of more specific elements. Aspects that can be documented are the editorial practise (**<editorialDecl>**), the project context in which the electronic text was realised (**<projectDesc>**), a declaration of all XML elements used in the encoding (**<tagsDecl>**), declaration of reference systems used in the encoding (**<refsDecl>**), and a declaration of any classification schemes used to classify

the text (**<classDecl>**). When the **verse** TEI module is included in the TEI schema, the system for metrical analysis can be declared in the **<metDecl>** element. When the TEI schema includes the **textcrit** TEI module for the encoding of text-critical variants, the system used for variant encoding may be documented in **<variantEncoding/>**.

3.3 The Profile Description

The profile description, in the **<profileDesc>** element, is the third major section of the TEI header. It can be used to document all kinds of non-bibliographic information about an electronic text, either as loose prose in (a) paragraph(s), or in minimally one of more specific elements. The most important subsections are:

- **<creation>**: information about the creation of a text
- **<langUsage>** (language usage): information about the languages used in the text
- **<textClass>** (text classification): classification of the contents of the text, according to a classification scheme

Besides these general subsections, some TEI modules add other specific subsections to the profile description. When the **transcr** TEI module for the description of primary sources is included in a TEI schema, following element can be used in **<profileDesc>**:

- **<handNotes>**: identification of the different hands in a primary document

10

3.3.1 Creation

If there are any details worth recording about the actual place or time of creation of the source text, this can be done in the **<creation>** element, as a loose prose description. This may be useful when the text was created long before its publication, as an exact situation in place and time can be important to certain types of research (e.g. study of (diachronic) linguistic variation). It is worth pointing out the difference between the circumstances about the creation of a text (which must be recorded in this **<creation>** element), and the bibliographic circumstances of its publication (which must be recorded in a bibliographic description in the **<sourceDesc>** subsection of the TEI header).

For example:

```
<creation>Original written in <date when="1831">1831</date> in <name type="city">Paris</name>.</creation>
```

.....

10 Besides these elements, the **corpus** TEI module can add more elements to describe specific aspects of the compilation of a language corpus. These are not covered in this introductory tutorial; for a full reference, see section [15.2 Contextual Information](#) of the TEI Guidelines.

Summary

The **<creation>** element can provide a prose description about particularities on the creation of a text, such as its actual time and place of writing.

3.3.2 Language Usage

The **<langUsage>** subsection of the profile description provides room to describe the different languages used in the text. Each language must be described in a distinct **<language>** element. It may contain a prose description of the language (or dialect), and must provide a formal identification code for this language in the **@ident** (identifier) attribute. When appropriate, the distribution of this language over the text contents can be stated as a percentage in a **@usage** attribute. Section [vi.1. Language identification](#) of the TEI Guidelines offer recommendations for the constructions of the formal language identification codes for the **@ident** attribute. It is important that these codes correspond to the values of the **@xml:lang** attributes to identify phrases in that language in the electronic text.

For example, the languages used in the English translation of *The Wild Ass's Skin* could be defined as follows:

```
<langUsage>
  <language ident="en" usage="98">English</language>
  <language ident="fr" usage="1">French</language>
  <language ident="ar" usage="1">Arabic</language>
</langUsage>
```

Summary

The languages used in an electronic text can be formally declared in the **<langUsage>** subsection of the profile description. Each language can be described in a separate **<language>** element, which must contain a formal identification code in the **@ident** attribute, and can provide details about the distribution of this language in the text in a **@usage** attribute.

3.3.3 Text Classification

The contents of the text can be classified according to one or more classification schemes in the **<textClass>** subsection of the profile description. It can be done by means of three specific elements:

- **<keywords>**: a list of keywords in a given keyword list
- **<classCode>**: a classification code in a given classification scheme

- **<catRef/>**: a list of pointers to specific categories in a given taxonomy

In general, these elements allow for two kinds of classification:

- reference to an external classification scheme, which uses either subject headings (**<keywords>**), or classification codes (**<classCode>**)
- reference to specific categories in an internally defined taxonomy (**<catRef/>**)

The **<keywords>** and **<classCode>** elements fulfil a similar role: they allow you to use classification categories defined in external classification schemes. If such a scheme defines categories in terms of subject headings, the **<keywords>** element should be used to refer to those keywords; if the scheme defines categories in terms of classification codes, the **<classCode>** element should be used. The classification scheme must be identified in the **@scheme** attribute, which contains a pointer to its declaration in a **<taxonomy>** element inside the **<classDecl>** subsection of the encoding description (see [3.2.5 The Classification Declaration \(p. 29\)](#)).

The **<keywords>** element must list the terms either in a series of **<term>** elements, or make use of a **<list>** structure. For example, *The Wild Ass's Skin* can be classified in terms of the Library of Congress Subject Headings scheme as follows:

```
<keywords scheme="#lcsch">
  <list>
    <item>Literature</item>
    <item>Fiction and juvenile belles lettres</item>
    <item>Literature--Translations into English</item>
  </list>
</keywords>
```

...and / or in terms of the Dewey Decimal Classification Scheme like this (see code 082 in the library record above):

```
<classCode scheme="#DDC">843.7</classCode>
```

When you have defined your own classification system in the encoding description (see [3.2.5 The Classification Declaration \(p. 29\)](#)), you can refer to one of its categories by means of the **<catRef/>** element. This is an empty element that must point to the category definitions in a **@target** attribute. This is basically a list of pointers to the **@xml:id** attributes of the relevant categories in one of the **<taxonomy>** elements you defined in the profile description section of the TEI header. If the reference to the category does not suffice, the **@scheme** attribute may point to the declaration of the relevant taxonomy containing the category. For example, *The Wild Ass's Skin* could be classified using the Balzac-specific classification scheme declared above as follows:

```
<catRef target="#BCS.phil" scheme="#BCS"/>
```

Summary

An electronic text can be classified in the **<textClass>** subsection of the profile description. A classification can use a keyword (**<keyword>**) or a classification code (**<classCode>**) defined in an external classification scheme. The **@scheme** attribute must be used to refer to the declaration of any external classification scheme in the **<classDecl>** subsection of the encoding description. Alternatively, the classification can be done using internally defined classification categories defined in the **<classDecl>** subsection of the encoding description. This is done by pointing to the definition of the relevant classification categories in the **@target** attribute of a **<catRef/>** element.

3.3.4 Document Hands

When the **transcr** TEI module for the transcription of primary sources is included in the TEI scheme, the profile description contains an additional element for the declaration of the different hands occurring in the document **<handNotes>**.

Crosslink

See

- TBE module [6: Primary Sources](#) for a discussion of the elements in the TEI **transcr** module
- TBE module [8: Customising TEI, ODD, Roma](#) for a discussion of how to include the TEI **transcr** module in a TEI schema

Each hand that occurs in the source text can be identified in a **<handNote>** element, containing a prose description of its characteristics. It should be identified with an **@xml:id** attribute, and can contain additional attributes for formalised documentation of the script or writing style (**@script**), writing medium (**@medium**), and an indication of the prominence of this hand in the text (**@scope**). If the document hand can be ascribed to a specific person, this person can be identified in the **@scribe** attribute.

For example, the **<handNotes>** element could be used to identify the hand in which the previous owner of the book has added some annotations (supposed we wanted to transcribe these as well), as well as the Arabic text in this example:

THE TALISMAN.

31

the surface impression, that for a moment he was not sure that he had cut anything away after all.

“The craftsmen of the Levant have secrets known only to themselves,” he said, half in vexation, as he eyed the characters of this Oriental sentence.

“Yes,” said the old man, “it is better to attribute it to man’s agency than to God’s.”

The mysterious words were thus arranged :—

لو ملكتنى ملكت آل الكر
ولكن عمرك ملكى
واراد الله هكذا
اطلب وستنال مطالبك
ولكن قسن مطالبك على عمرك
وهي هاهنا
فبكل مرامك استسنزل ايامك
أتريدنى
الله يجيبك
آمين

Or, as it runs in English :—

POSSESSING ME THOU SHALT POSSESS ALL THINGS.
BUT THY LIFE IS MINE, FOR GOD HAS SO WILLED IT.
WISH, AND THY WISHES SHALL BE FULFILLED ;
BUT MEASURE THY DESIRES, ACCORDING
TO THE LIFE THAT IS IN THEE.
THIS IS THY LIFE,
WITH EACH WISH I MUST SHRINK
EVEN AS THY OWN DAYS.
WILT THOU HAVE ME? TAKE ME.
GOD WILL HEARKEN UNTO THEE.
SO BE IT!

Fig. 3.

```
<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>
```

Summary

When the **transcr** TEI module for the transcription of primary sources is included in the TEI scheme, different hands in the source text can be identified in **<handNote>** elements inside a **<handNotes>** subsection of the profile description.

3.3.5 Summary

With the information about the text's creation, languages, and classification in place, the **<profileDesc>** section of the TEI header for our sample text could look as follows:

```

<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="#lcsb">
      <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>

```

Summary

The profile description (**<profileDesc>**) is the third section of the TEI header. It describes all kinds of non-bibliographic information about an electronic text, either as loose prose in (a) paragraph(s), or in minimally one of more specific elements. Aspects relating to the creation of the text can be documented in **<creation>**, the languages used in the document can be declared in **<langUsage>**, and a text classification can be provided in **<textClass>**. When the **transcr** TEI module for the representation of primary sources is included in the TEI schema, the different hands occurring in the source text can be formally documented in a **<handNotes>** element.

3.4 The Revision Description

The fourth and final part of the TEI header is reserved for a detailed record of the revisions that have been made to the electronic text, in **<revisionDesc>**. Each revision is described in a dedicated **<change>** element. Additionally, it makes sense to formally identify the exact date of

the change in a **@when** attribute, and the person responsible for the change in a **@who** attribute. The latter points to the definition of a person responsible for some aspect of the electronic text, which is probably defined in the **<titleStmt>** subsection of the file description section of the TEI header.

Although ordering is arbitrary, it makes sense to rank the changes in chronological order, either ascending or descending. This optimises both readability and maintainability of this logbook, so that it can provide an instant overview of the complete history of the electronic text. For example:

```
<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>
```

Summary

The complete revision history of an electronic text can be documented in the **<revisionDesc>** section of the TEI header. Each change to the electronic file can be categorised and recorded in a separate **<change>** element. The **@when** attribute can record the date of change, while the **@who** attribute can be used to refer to an identified person responsible for some aspects of the text.

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 exemplar may be in loan, you find another reference to *The Wild Ass's Skin* in following record (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. English		
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 -- tGaussart 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 tlily of the valley -- gv. 28. The tillustrious Gaussart -- gThe tmuse of the department -- gv. 29. The told maid -- gThe tcabinet of antiquities -- gv. 30-32. tLost illusions g: The two 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 treperory of La comédie humaine of Honoré de Balzac -- gv.52-53. Les tcontes drolatiques.		
600 1 0	aBalzac, Honoré de, d1799-1850.		

Fig. 4.

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 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 the individual TEI headers.

Following example gives an impression of what the TEI header for an electronic edition of *La Comédie Humaine* might look like:

```

<teiCorpus>
  <!-- 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 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>
            <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 (CTB)</name>
          <name type="institution">Royal Academy of Dutch Language and Literature</name>
          <street>Koningstraat 18</street>
          <postCode>9000</postCode>
          <name type="city">Gent</name>
          <name type="country">Belgium</name>
        </address>

```

```

<idno type="ISBN">0-00-000000-9</idno>
<availability status="free">
  <p>Published under a <ref target="http://creativecommons.org/licenses/by-sa/3.0/">Creative
    Commons Attribution ShareAlike 3.0 License</ref>.</p>
</availability>
<date when="2110-01-01">1 January 2110</date>
</publicationStmt>
<seriesStmt>
  <title>The TBE collection: sample texts encoded with TEI.</title>
  <idno type="ISSN">0000-0001</idno>
  <respStmt>
    <name>Edward Vanhoutte</name>
    <resp>compiler</resp>
  </respStmt>
</seriesStmt>
<notesStmt>
  <note>OCR scanning done at KANTL, Gent.</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>
  </editorialDecl>
  <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 TEI</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="lsh">
    <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 -->
  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title>The Wild Ass's Skin: an electronic edition</title>
        <editor role="translator" xml:id="EM">Ellen Marriage</editor>
        <editor role="preface" xml:id="GS">George Saintsbury</editor>
      </titleStmt>
      <editionStmt>
        <edition n="2.0">
          <title>Version 2.0, enriched with thematic annotations.</title>
          <date when="2010">2010</date>
        </edition>
        <respStmt>
          <name>Melissa Terras</name>
          <resp>Added thematic annotations.</resp>
        </respStmt>
      </editionStmt>
      <extent>572 Kb</extent>
      <publicationStmt>
        <idno type="ISBN">0-00-000000-0</idno>
        <date when="2010-01-01">1 January 2010</date>
      </publicationStmt>
      <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>288 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="#lcsb">
            <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>
```

11

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.

5. Summary

After this overview of the most current header sections, it is time to put them all together and illustrate how a fairly detailed header for our sample text could look:

.....

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

```

<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>The Wild Ass's Skin: an electronic edition</title>
      <author xml:id="HdB">Honoré de Balzac</author>
      <editor role="translator" xml:id="EM">Ellen Marriage</editor>
      <editor role="editor" xml:id="TBEcrew">The TBE crew</editor>
      <editor role="preface" xml:id="GS">George Saintsbury</editor>
    </titleStmt>
    <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 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>
        <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>
  <editionStmt>
    <edition n="2.0">
      <title>Version 2.0, enriched with thematic annotations.</title>
      <date when="2010">2010</date>
    </edition>
    <respStmt>
      <name>Melissa Terras</name>
      <resp>Added thematic annotations.</resp>
    </respStmt>
  </editionStmt>
  <extent>572 Kb</extent>
  <publicationStmt>
    <publisher>Centre for Scholarly Editing and Document Studies (CTB)</publisher>
  </publicationStmt>

```

```

<istributor>Centre for Computing in the Humanities (CCH) - King's College London</istributor>
<pubPlace>Gent</pubPlace>
<address>
  <name type="institution">Centre for Scholarly Editing and Document Studies (CTB)</name>
  <name type="institution">Royal Academy of Dutch Language and Literature</name>
  <street>Koningstraat 18</street>
  <postCode>9000</postCode>
  <name type="city">Gent</name>
  <name type="country">Belgium</name>
</address>
<idno type="ISBN">0-00-000000-0</idno>
<availability status="free">
  <p>Published under a <ref target="http://creativecommons.org/licenses/by-sa/3.0/">Creative Commons
  Attribution ShareAlike 3.0 License</ref>.</p>
</availability>
<date when="2010-01-01">1 January 2010</date>
</publicationStmt>
<seriesStmt>
  <title>The TBE collection: sample texts encoded with TEI.</title>
  <idno type="ISSN">0000-0001</idno>
  <idno type="installment">1</idno>
  <respStmt>
    <name>Edward Vanhoutte</name>
    <resp>compiler</resp>
  </respStmt>
</seriesStmt>
<notesStmt>
  <note>OCR scanning done at KANTL, Gent.</note>
</notesStmt>
<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>288 p.</extent>. Translated by <editor role="translator">Ellen
  Marriage</editor> ; preface by <editor role="preface">George Saintsbury</editor>.</bibl>
</sourceDesc>
</fileDesc>
<encodingDesc>
  <editorialDecl>
    <correction method="markup">
      <p>Apparent errors have been corrected using the sic / <corr> elements, wrapped in a <choice>
      element.</p>
    </correction>
  </editorialDecl>

```

```

<normalization method="markup" source="http://www.oed.com/">
  <p>Spelling has been modernised using the <orig> / <reg> elements, wrapped in a <choice>
  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 TEI</soCalled>, aiming
  at providing a collection of prime exemplar TEI encoded materials.</p>
</projectDesc>
<tagsDecl>
  <rendition scheme="css" xml:id="division">display:block; margin: 1em;</rendition>
  <rendition scheme="css" xml:id="paragraph">display:block; margin-bottom: 0.5em;</rendition>
  <rendition scheme="css" xml:id="red">color:red;</rendition>
  <rendition scheme="css" xml:id="smallcaps">font-variant:small-caps;</rendition>
  <rendition scheme="css" xml:id="large">font-size:large;</rendition>
  <!-- ... -->
  <namespace name="http://www.tei-c.org/ns/1.0">
    <tagUsage gi="div" occurs="300" withId="300" render="#division">Marks text divisions in the source
    text.</tagUsage>
    <tagUsage gi="p" occurs="8302" withId="8300" render="#paragraph">Marks paragraphs in the source
    text.</tagUsage>
    <!-- ... -->
  </namespace>
</tagsDecl>
<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.

```

</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="lsh">
    <bibl>
      <title>Library of Congress Subject Headings</title>
    </bibl>
  </taxonomy>
  <taxonomy xml:id="BCS">
    <category xml:id="BCS.man">
      <catDesc>Studies of Manners</catDesc>
      <category xml:id="BCS.man.priv">
        <catDesc>Scenes from Private Life</catDesc>
      </category>
      <category xml:id="BCS.man.prov">
        <catDesc>Scenes from provincial life</catDesc>
        <category xml:id="BCS.man.prov.cel">
          <catDesc>The Celibates</catDesc>
        </category>
        <category xml:id="BCS.man.prov.par">
          <catDesc>Parisians in the Country</catDesc>
        </category>
        <category xml:id="BCS.man.prov.jeal">
          <catDesc>The Jealousies of a Country Town</catDesc>
        </category>
      </category>
      <category xml:id="BCS.man.par">
        <catDesc>Scenes from Parisian life</catDesc>
        <category xml:id="BCS.man.thir">
          <catDesc>The Thirteen</catDesc>
        </category>
        <category xml:id="BCS.man.rel">
          <catDesc>Poor Relations</catDesc>
        </category>
      </category>
    </category>
  </taxonomy>

```

```

</category>
<category xml:id="BCS.man.pol">
  <catDesc>Scenes from political life</catDesc>
</category>
<category xml:id="BCS.man.mil">
  <catDesc>Scenes from military life</catDesc>
</category>
<category xml:id="BCS.man.cou">
  <catDesc>Scenes from country life</catDesc>
</category>
</category>
<category xml:id="BCS.phil">
  <catDesc>Philosophical studies</catDesc>
</category>
<category xml:id="BCS.ana">
  <catDesc>Analytical studies</catDesc>
</category>
</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>
  <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>
  </textClass>

```

```

</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>

```

6. What's next?

You have reached the end of this tutorial module covering prose markup with TEI. You can now either

- proceed with [other TEI by Example modules](#)
- have a look at the [examples section](#) for the prose module.
- take an interactive test. This comes in the form of a set of multiple choice questions, each providing a number of possible answers. Throughout the quiz, your score is recorded and feedback is offered about right *and* wrong choices. Can you score 100%? Test it [here](#)!