TEI BY EXAMPLE

MODULE 0: INTRODUCTION TO TEXT ENCODING AND THE TEI

Ron Van den Branden

Edward Vanhoutte

Melissa Terras

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

Last updated September 2020

Licensed under a Creative Commons Attribution ShareAlike 3.0 License

TABLE OF CONTENTS

4. COCOA..... 1

4. COCOA

In the next example, the sample text is encoded in COCOA. This encoding scheme shares with the LaTeX example above its non-XML character, but differs in that COCOA is a "descriptive" markup scheme. It provides a simple means to distinguish user-defined categories in a text, by labeling them unambiguously by means of one-letter tag names. There are two possibilities: either the text is encoded in the tag (e.g., `<H Review>` identifies the text "Review" as belonging to the category "H" (for "heading")), or a tag is numbered (e.g., `<P 1>` indicates that the text following it is part of the first paragraph). This enables the encoder not only to distinguish all text structures (heading ("H"), paragraph ("P"), footnote ("F"); but also to distinguish between the different textual phenomena that occur as italicised text (book title ("B"), emphasis ("E"), term ("T")). Moreover, the typographically unmarked proper name "Goethe" can be tagged as such as well ("N").

```
<H Review>
 <P 1><B Die Leiden des jungen Werther>&lt;F 1>by <N Goethe > is an
 <E exceptionally> good example of a book full of <T Weltschmerz>
```

Example 3. A COCOA example.