

Villanova's Digital Library Standards for Transcription 2.3 (2020-03-12)

Many documents in the digital library come from archival collections which contain letters, notes, telegrams, etc. One of the goals of the digital library is to transcribe these handwritten documents so that they are more easily searchable and accessible to the public. Transcription involves copying every letter, notation, and bit of punctuation into a machine readable format, for us, .doc files which will later be converted to .pdf files.

Villanova's Digital Library Standards for Transcription are as follows:

1. Copy every word of a document in to a .doc file (**do not** use .docx).
2. Each file should be named consistent with the following guidelines regardless of document type (i.e. letter, telegram, form, etc). This file name includes the PID with the .doc file extension, thus PID.doc (e.g. **vudl:6780.doc**). The PID can be located in the URL for the specific resource. (E.g. in this URL the PID is in bold: <https://digital.library.villanova.edu/Item/vudl:6780> .
3. Please include a header and footer in your transcription. The header should include, "Title:", the type of document "Letter", the author and recipient, and date, all as reflected in the "Title" of the metadata record.
 - **Header:** Letter, A.M. Thackara and General Sherman, July 30, 1881.The footer should follow the following format "Transcribed by: Your Name, the date you completed the transcription"
 - **Footer:** Transcribed by: Teri Ann Incrovato 9/17/07.
4. Keep the original writer's errors (spelling, punctuation, capitalization, etc.)
5. If you come to a word that is illegible mark it by [...] and include an endnote if you think you can make out some of the letters of the word AND/OR if there is more than one word missing, in which case, note how many words are missing.
6. For words that you are uncertain should be noted by an endnote (i.e. "Best guess").
7. Strikeouts should be noted using the 'strikethrough' option in the Format > Font > Effects section, check the "strikethrough" box. (~~For Example~~)
8. Ink blots or slips of the pen may be omitted unless they obscure the meaning of a word then they should be noted in brackets near where they occur.
9. The original formatting of the document should be adhered to as much as possible (e.g. placement of date, salutation, page, etc.) Line breaks will not be kept, while page breaks will be noted in the text:

For example:
Last line of text p. 1
(space)
[p.2]
(space)
First line of text p. 2
10. Words that were inserted into the text using a "^" or arrow should be included in the transcription but in <insertion: word >.
11. Emphasis should be kept where possible, underlined words should be underlined.
12. Superscripted letters should be included as ^{superscripted} in the text.
13. If more than one person wrote on the document a distinction should be made in the text (e.g. the primary writer in normal font, the second in italics). If both authors are known then a note should be added explaining who each author is.
14. In the cases were words have been shortened (i.e. "yr. most humble and obedt. servant") keep the original author's text as is. An additional example, Sept. remains Sept. not September. Likewise, symbols should be kept as closely as possible, a "plus sign" indicative of "and" is kept as + and not transliterated.

15. If your letter has letterhead, transcribe the machine printed letterhead but use THE SMALL CAPS FEATURE found under Format > Font > Effects > Small Caps. The small caps feature is still case sensitive so you can still show capital letters vs. lower case letters while using this special feature.
16. A note on Forms & Telegrams: For forms or anything with machine printed text use THE SMALL CAPS FEATURE found under Format > Font > Effects > Small Caps to differentiate printed matter from handwritten matter. Handwritten text on a form should be typed in “normal” but *italicized* font. If there are “blanks” on a form like this: *Bobby* you can create the front and end of the line by using an underscore “_” then underline the written in word by highlighting the word while holding down Control + U.
17. Endnotes : Please use endnotes not footnotes (Insert > Reference > Footnotes). Please be sure that Arabic numbers are selected under “number format” (1, 2, 3 ...). Use these common endnotes when possible.

Common Endnotes:

- Indecipherable word
- Best guess
- Partially overwritten word
- Best guess - written diagonally on the page in a different hand.
- Italicized segment written over [top left] part of letter, perpendicular to the main body of letter, by author.
- Name authorities: i.e.
 - Sherman, Thomas Ewing, 1856-1933.
 - Sherman, Mary Elizabeth, 1852-1925.
 -

**Please use names as found in standard lists if one exists for that collection (or in National Authority databases such as the Library of Congress Name Authority:
<https://authorities.loc.gov/webvoy.htm>)**

Special Considerations

1. The docketing, that is, the information written at the head of the letters in pencil by an earlier archivist to identify them should be noted by: [...] with an endnote that reads: Docketed: “whatever is written in that note”. See also the example letter.
2. Letters in some collection were written “out of order” that is to say after the first page the letter continues on the “third” page of the document then the author returns to the back of the first page. In these instances the flow of the letter should not be interrupted, but the sequence of pages should be listed in the transcription as the document appears to be written.
 For example:
 [p. 1]
 The text should be able to be read as a continuation of the sentence from the first
 [p. 3] ← Add endnote here that reads: “Author wrote the letter in the following order: p.1, 3, 2, 4.”
 to the second page with ONLY the bracketed number as an indication that the text is in a different sequence from a traditional front, back, front, back ordering of the letter. Further more an endnote should be added, see above.
3. Blank pages should be noted in the transcriptions
 For example
 [p. 1]
 Text of letter here
 [p. 2] ← endnote that reads, “blank”
 [p. 3] ← endnote that reads, “blank”
4. When letters are “continued” from the back of the last page on to the front of the 1st page of the letter (perpendicular to the text) they should be added to the end of the transcription for ease of reading. The page number should read as [p. 1 cont.] and the standard endnote should be applied to the first word of this section.

REFERENCE

Michael E. Stevens and Steven R Burg. *Editing Historical Documents: A Handbook of Practice*. Altamira Press, 1997.