The Digital Ælfric Project: Rules for Encoding

Dot Porter, Indiana University

Rev. Sept. 2010

A Note on Terminology: "Elements" are the main descriptors in XML encoding, and are enclosed in pointed brackets. "Attributes" occur inside elements to further describe the content of the elements. Attributes contain values, while element contain textual content.

```
<element attribute="value">CONTENT</element>
```

In narrative description, elements are presented in pointed brackets. "Here, we are talking about `<element>`." Attributes are identified as being preceded by an "@" symbol. "Here, we are talking about `@attribute`."

1.1. Initial markup

1.1.1. TEI Header & Witness List

Each homily is first encoded in in a basic TEI P5 wrapper. This wrapper includes a TEI Header that has three main sections used by the project:

File Description: bibliography

*File Description (<fileDesc>): This section includes title information (including the Short Title and Cameron Number for the individual homilies), a brief publication statement, and other metadata about the source text. It also includes a responsibility statement, which lists all people responsible for the encoded text (editors, taggers) and each individual is assigned a unique ID that is used in the encoding to show who is responsible for various tags:

```
<fileDesc>
<titleStmt>
<title type="short-title"> ÆCHom I, 15 </title>
<title type="cameron-number" xml:id="B1.1.17">B1.1.17</title>
<title>Easter</title>
<editor>Electronic Aelfric Project</editor>
<respStmt>
<resp>Editor</resp>
<name xml:id="AK">Aaron Kleist</name>
</respStmt>
<respStmt>
<resp>Research Associate</resp>
<name xml:id="DP">Dorothy Porter</name>
</respStmt>
<respStmt>
<resp>Tagging</resp>
<name xml:id="WDW">William Daniel Whittaker</name>
</respStmt>
</titleStmt>
<publicationStmt>
<p>Text from Clemoes 1955-6, 219-32, provided by the Dictionary of Old
English
Corpus. Variants added by the Electronic Aelfric project. Citation: Cited by
page and line no. from Thorpe 1844-6, I, 220-28. Biblical passages not
printed
by Thorpe are identified by their book and chapter no. in square
brackets.</p>
</publicationStmt>
<sourceDesc>
<p> Homily by Aelfric of Eynsham. Various manuscripts to be identified.</p>
<p>U: 3-18</p>
</sourceDesc>
</fileDesc>
```

Encoding Description: rules for encoding and abbreviation descriptions

Encoding Description (`<encodingDesc>`): This section describes the General Rules for encoding (which are also described in more detail on this page), and provides a list of abbreviations and their descriptions. As in the responsibility statement, each abbreviation is given a unique ID (@xml:id) which can then be referenced in the document encoding (for complete instructions on encoding abbreviations see --):

```
<p>Abbreviation List: Old English
<list>
<item>General Abbreviations
<list>
<item xml:id="macron-M">Suspended "m" or "n", indicated by a macron over
the "M"</item>
<item xml:id="macron-A">Suspended "m" or "n", indicated by a macron over
the "A"</item>
<item xml:id="macron-E">Suspended "m" or "n", indicated by a macron over
the "E"</item>
<item xml:id="macron-U">Suspended "m" or "n", indicated by a macron over
the "U"</item>
<item xml:id="macron-I">Suspended "m" or "n", indicated by a macron over
the "I"</item>
```

```
</list>
</item>
<item>Specific Abbreviations: suspensions
<list>
<item xml:id="cross-thorn">"þæt" abbreviated by a cross through the
ascender of the "þ"</item>
<item xml:id="þonne">"þonne" abbreviated by "þon" with a macron over the
"O" and/or the "N"</item>
<item xml:id="macron-wynn">"cwæð" abbreviated by "c-wynn" with a macron
over the "wynn"</item>
</list>
</item>
</list>
</p>
```

Revision Description: Noting Changes over Time

Revision Description (<revisionDesc>): This section is where all taggers will note changes they make to the XML document as encoding progresses. At least once-a-week (more if it is found useful), add a new <change> element inside the <revisionDesc>, indicate the date (in @when) and who you are (place the @xml:id assigned to your name in the File Description as the value of @who), and write a short description of what you've done (how many lines, what markup precisely) in a <p> tag:

```
<revisionDesc>
<change when="2006-06-08" who="#DP">
Created initial XML file. Copied header information from .doc file provided
by
<name corresp="#AK">Aaron Kleist</name>. Divided text into segments and
assigned @n values according to the numbering provided in the original file.
</change>
<change when="2006-06-09" who="#DP">
Added witness list. Added short title and cameron number to titleStmt.
</change>
<change when="2006-07-05" who="#DP">
Initial markup at the beginning of the homily, first 2 segments.<
</change>
<change when="2006-07-11" who="#WDW">
Day Two.
-I was just thinking... could you leave certain notes? especially the
ones that I can change once you explain it to me how you want it done. that
way
I can find them later. I will delete the notes after they are corrected
for...
or however works best.
<hi rend="dot">Sure. I've tried to leave in the more interesting notes,
though I did
cut out many of the punctuation ones after I added the markup. I answered
many of
your questions, and indicated too when it's okay for you to delete the notes
if you
want to. I won't in case you want them. Definitely don't delete any of the
biblical citation notes!
</hi>
-How important is spacing within the xml code itself. I mean the spacing
before and after 'app,' etc.
<hi rend="dot">It's important. Try not to add extra spaces.</hi>
```

```
-It may be easier to show you what I mean than trying to explain in text.
-Two hours.
</change>
<change when="2006-07-13" who="#WDW">
Third day.
-Revised capitalization, the last two examples show on the wiki for
the 'Variant' section (which I had not read thoroughly before) and
'Abbreviations with Variants' section (but I do not understand what is going
on
with that one) for page one.
<hi rend="dot">Actually, there is an example of "Abbreviations within
Variants" in your work.
Look down in the XML to lines 509 and 510 (if the numbering isn't the same,
search for #DP).
This shows an abbreviation nested in a variant. Let me know if you still have
questions
about this.</hi>
-Several questions/references in notes regarding punctuation.
<hi rend="dot">I think I got them all. From here on out, just mark
variants in punctuation as you would any other variant - only the content of
the rdg tags will be periods and colons instead of complete words or
phrases.</hi>
-How should words repeated in the manuscript be referenced? (see note: pg 4-
line 11)
<hi rend="dot">What you did on p. 4 is great, and it's what should be done
whenever something
appears in one source but not in the other - put in an app with only one
rdg.</hi>
-How do we handle words that are not listed on the DOE but in the manscript?
a similar question to the one
directly above. (see note pg 2-line 13)
<hi rend="dot">Exactly the same. Just one app/rdg set.</hi>
-Started on page two. -Hour and a half.
</change>
</revisionDesc>
```

Witness List: List of all Manuscripts

Following the TEI Header but before the homily text comes the witness list. This lists all the manuscripts in which the eAelfric homilies are found plus a witness listing for the Dictionary of Old English Corpus (DOE) (the homily transcriptions used as base for this project are from the DOE, so we need a DOE witness to record variants and abbreviation expansions from the transcription). Each individual witness is assigned a unique ID. To simplify encoding, all witnesses except the DOE and the witness currently being encoded is commented out. In the short example below, only the DOE and D witnesses are available; all other witnesses are commented out using the syntax <!-- xxx -->:

```
<listWit n="manuscripts">
<witness xml:id="DOE">Dictionary of Old English Corpus</witness>
<!--<witness xml:id="A">London, British Library, Royal 7 c. xii</witness>
<witness xml:id="B">Oxford, Bodleian Library, Bodley 343</witness>
<witness xml:id="C">Cambridge, Corpus Christi College 303</witness>-->
<witness xml:id="D">Oxford, Bodleian Library, Bodley 340 and Oxford, Bodleian
Library, Bodley 342</witness>
<!--<witness xml:id="E">Cambridge, Corpus Christi College 198</witness>
<witness xml:id="F">Cambridge, Corpus Christi College 162</witness>
<witness xml:id="fB">Cambridge, Jesus College 15</witness>-->
```

```
</listWit>
```

1.1.2. TEI body, seg

After the Witness List comes the text of the homily. The entire homily text is nested in the TEI tags <body> and <p>, and the text is divided into segments using the TEI tag <seg>.
Segment divisions and segment numbers are as provided by the DOE, based on page and line numbering from Thorpe's edition:

```
<body>
<p>
...
<seg n="220.5">Ða ða crist bebiriged wæs. þa cwædon þa iudeiscan. to heora
ealdermen
pilate;</seg>


<seg n="220.6">La leof se swica þe her ofslegen is cwæð gelomlice þa ða he on
life
wæs. þæt he wolde arisan of deaðe on ðam þriddan dæge;</seg>
...
</p>
</body>
```

As the project continues, we will want to add the numbering from Clemoes' published edition. When we do this, we will need to make several decisions:

1. Whether to *replace* the DOE/dissertation numbering with the edition numbering,

```
<seg n="222.4">Ða ða crist bebiriged wæs. þa cwædon þa iudeiscan. to heora
ealdermen
pilate;</seg>
<seg n="222.5">La leof se swica þe her ofslegen is cwæð gelomlice þa ða he on
life
wæs. þæt he wolde arisan of deaðe on ðam þriddan dæge;</seg>
```

2. OR to *add a new attribute* to hold the edition numbering,

```
<seg n="220.5" ed-n="222.4">Ða ða crist bebiriged wæs. þa cwædon þa
iudeiscan. to heora ealdermen pilate;</seg>
<seg n="220.6" ed-n="222.5">La leof se swica þe her ofslegen is cwæð
gelomlice þa ða he on life
wæs. þæt he wolde arisan of deaðe on ðam þriddan dæge;</seg>
```

3.OR to *replace* the DOE/dissertation numbering with the edition numbering
'and' add a new attribute to hold the DOE/dissertation numbering

```
<seg n="222.4" diss-n="220.5">Ða ða crist bebiriged wæs. þa cwædon þa
iudeiscan. to heora
ealdermen pilate;</seg>
<seg n="222.5" diss-n="220.6">La leof se swica þe her ofslegen is cwæð
gelomlice þa ða he on
    life wæs. þæt he wolde arisan of deaðe on ðam þriddan dæge;</seg>
```

1.2. Navigational markup

We need very basic markup to show the page breaks and line breaks in the manuscript. Use the <pb> (for page break) and <lb> (for line break) tags. Both of these tags take two attributes: @n (which records the number of the line or page) and @ed (which records the witness of the manuscript as recorded in the witness list). Page numbers should be the accepted folio numbering or pagination ("67v" for foliation, "104" for pagination).

'Line breaks and page breaks should come before, not after, the line or page being numbered.'
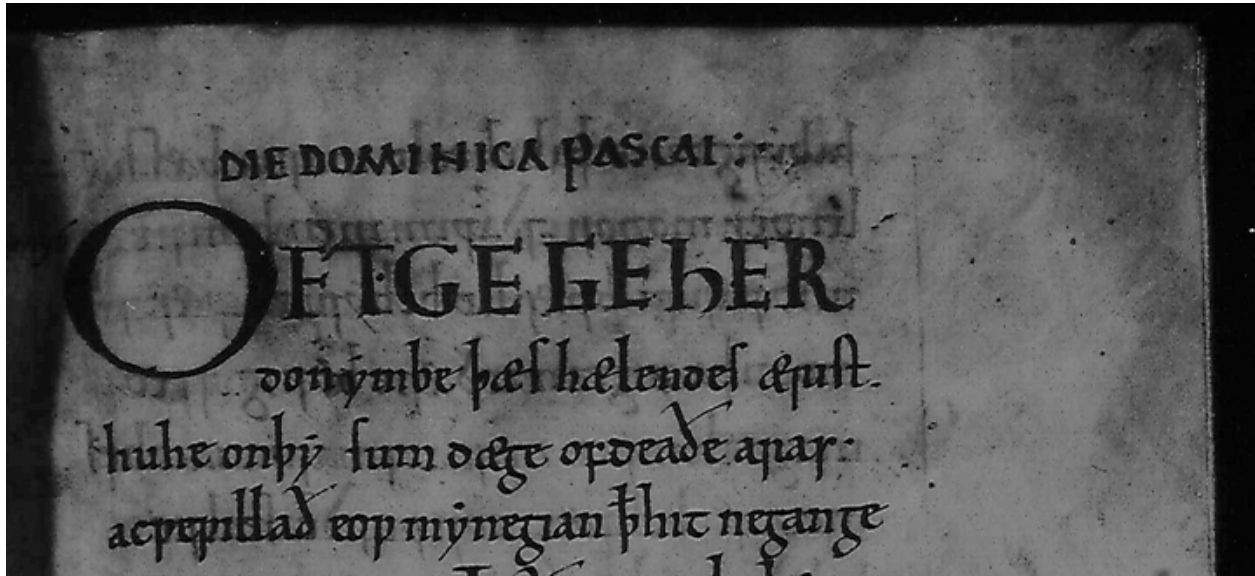

```
<pb ed="#U" n="3"/>
<lb ed="#U" n="1"/>dominica pasce et reliqua;
<lb ed="#U" n="2"/>Oft ge gehyr
<lb ed="#U" n="3"/>don ymbe ðæs hælendes ærist.
<lb ed="#U" n="4"/>hu he on þisum dæge of deaðe aras:
<lb ed="#U" n="5"/>ac we wyllað eow myningean. þæt hit ne gange
```

If your homily begins in the middle or the end of a folio, you should still place a <pb/> at the start of your file. If your homily begins in the middle of a line (as in the following example), mark the line number in a <lb/> but do not transcribe the text from the first part of the line.

If a line break comes in the middle of a word, place it right in the middle of the word! Do not add any extra spaces. See the break between line 2 and line 3 in the example above. Here it is in another view:

```
<lb ed="#U" n="2"/>Oft ge gehyr<lb ed="#U" n="3"/>don ymbe ðæs hælendes
ærist.
```

Line numbers are based on the number of lines on a folio. If a homily begins in the middle of a page you will need to count down from the top of the folio to find the starting line number.
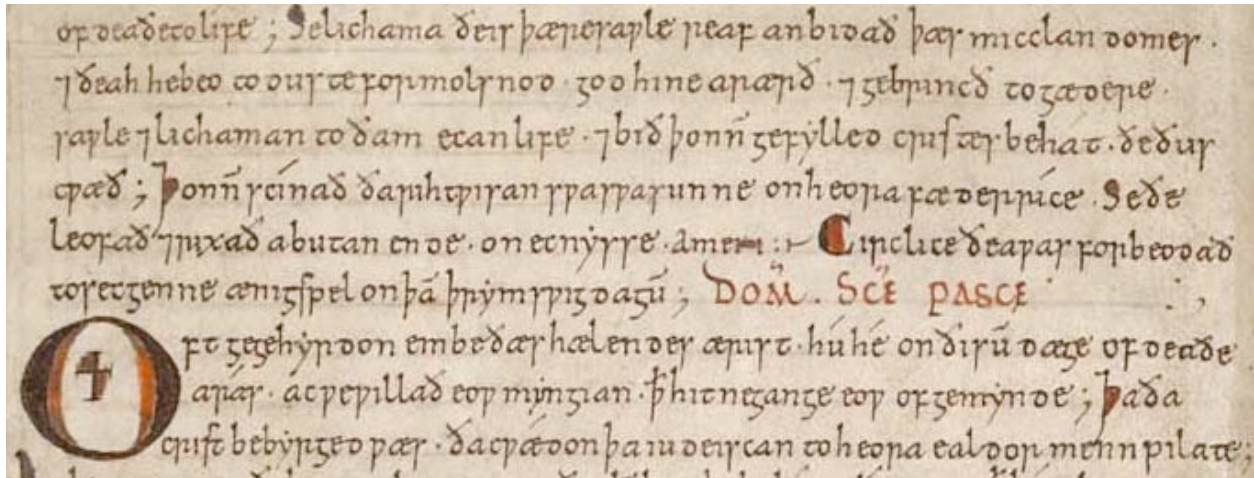

```
<pb ed="#K" n="47v"/>
<lb ed="#K" n="6"/>dom sce pasce
<lb ed="#K" n="7"/>Oft ge gehyrdon embe ðæs hælendes ærist...
```
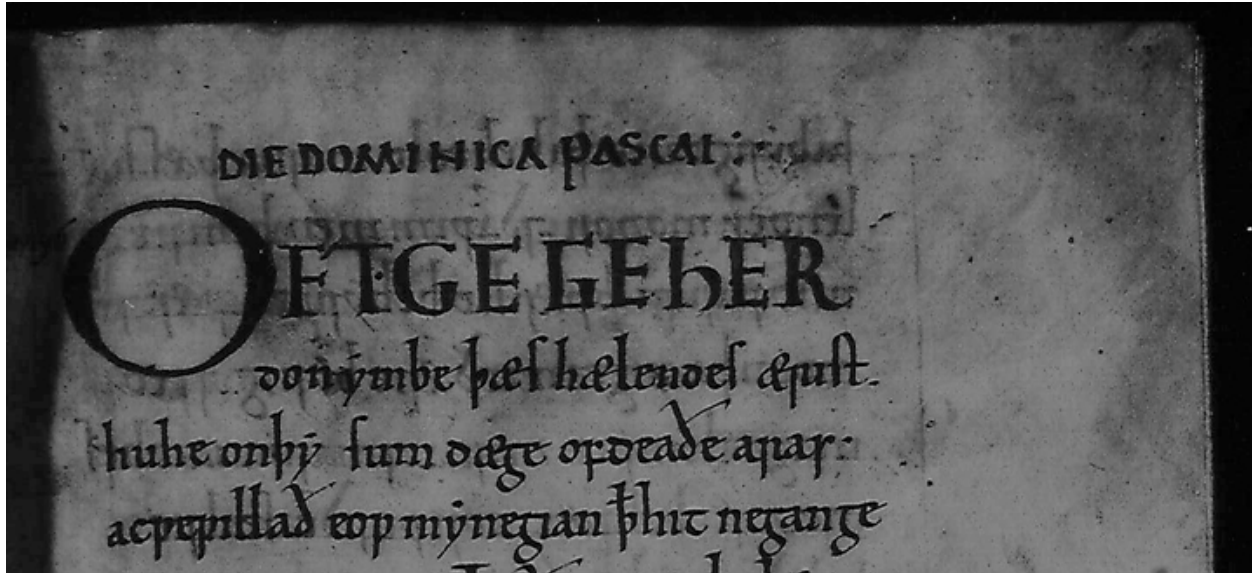
1.3. Variants

Using nested <app> and <rdg> tags, note all variants between the transcript and the manuscript. We are noting '*all variants*', including:

- variant words or phrasing (*dominica pasce Maria magdalene et maria iacobi.* vs. *DIE DOMINICA Pascai*) (also a capitalization variation!)
- variations in word order (*dæge of deaðe* vs. *deaðe of dæge*)
- spelling variations (*gehyrdon* vs. *geherdon*; *ðæs* vs. *þæs*)
- capitalization (*gehyrdon* vs. *Gehyrdon*)
- punctuation (*gehyrdon.* vs. *gehyrdon,* vs. *gehyrdon*)

For the most part, variants should be noted on the word or phrase level. For example, if a word begins with þ in the transcript but with ð in the manuscript, mark the whole word not the individual letter. The exception is with punctuation marks, which should be marked individually, unless of course they are part of a variation of phase.

Identify all <rdg> tags using the @wit attribute. @wit identifies the source for the variant according to the value assigned to the witness in the witness list (either the transcript - #DOE - or with # plus the @xml:id of the manuscript as assigned in the witness list). Any text that is the same in both the transcript and the manuscript should be left as-is, unmarked and undeleted.

```
<app>
<rdg wit="#DOE">dominica pasce Maria magdalene et maria iacobi. et
reliqua;</rdg>
<rdg wit="#U">DIE DOMINICA Pascai</rdg>
</app>
</seg>

Oft ge <app>
<rdg wit="#DOE">gehyrdon</rdg>
<rdg wit="#U">geherdon</rdg>
</app> ymbe <app>
<rdg wit="#DOE">ðæs</rdg>
<rdg wit="#U">þæs</rdg>
</app> hælendes ærist. hu he on <app>
<rdg wit="#DOE">þisum</rdg>
<rdg wit="#U">þysum</rdg>
</app> dæge of deaðe aras
<app>
<rdg wit="#DOE">:</rdg>
<rdg wit="#U">.</rdg>
</app> ac we wyllað eow myningean. þ hit ne
gange eow of gemynde;
```

1.3.1. Variants in Punctuation

Mark variants in punctuation as you mark other variants:

```
<app>
<rdg wit="#DOE">:</rdg>
<rdg wit="#U">.</rdg>
</app>
```

If there is punctuation present in the manuscript that is not in the edition, or vice versa, mark the single punctuation mark with its own app/rdg set:

```
<app>
<rdg wit="#U">.</rdg>
</app>
```

1.3.2. Variants in Capitalization

When the manuscript has a capital letter (usually at the beginning of a word), while the edition has a lower-case - or vice versa - these should be recorded as variant readings:

```
<app>
<rdg wit="#U">gehyrdon</rdg>
<rdg wit="#DOE">Gehyrdon</rdg>
</app>
```

In some cases, the manuscript may have characters that are clearly intended as upper-case (they are large-sized, rubricated or filled, and/or occur following punctuation) but are lower-case in appearance. See this typical example from #K:
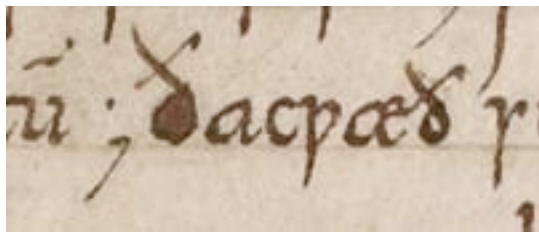
In this case, we will represent the character in the text with an upper-case glyph, with an alternative as an attribute value (to help with later browser view):

```
<c type="small" alt="ð">Ð</c>
```

where "small" on @type indicates this is in the form of a small eth and "ð" on @alt gives an "alternative" for rendering.

1.3.3. Variants of Missing Text

There are instances where sections of text (words, phrases, or longer spans) appear in the manuscript but not in the transcription, or in the transcription but not in the manuscript. This means that the transcription is using a reading from a different manuscript, and the text from the manuscript being edited was not included in the "main text" for some reason. We do want to note these variants.

In this case, the text should be placed in a single <rdg> with the content identified:

```
dæge of deaðe aras:
<app>
<rdg wit="#U">
ac we wyllað eow myningean.
</rdg>
</app>
þ hit ne
```

You may also come across instances where sections of text appear in both sources, but in slightly different places. In this case, the text should be marked as separate readings in both locations (using a modern English example):

```
DOE: Holy cow! Doug just ate an i-pod!
MS: Doug just ate an i-pod! Holy cow!
```

Marked as:

```
<app><rdg wit="#DOE">Holy cow!</rdg></app>
Doug just ate an ipod!
<app><rdg wit="#K">Holy cow!</rdg></app>
```

1.3.4. Finding the Extent of Variants

"I've reached a point where the transcript and the manuscript no longer correlate. There appears to be a substantial amount of text in one that does not appear in the other. Is there a way to figure out where the shared text picks up again?"

Example from Homily 16:

DOE has this:

```
ofteoð þa forgifenysse: þam bið oftogen; þa næs thomas þær an þæra
twelfa apostola; ða þomas com. þa cwædon his geferan him to: we gesawon
þone hælend;
```

London, British Library, Cotton Faustina A. ix looks like this:
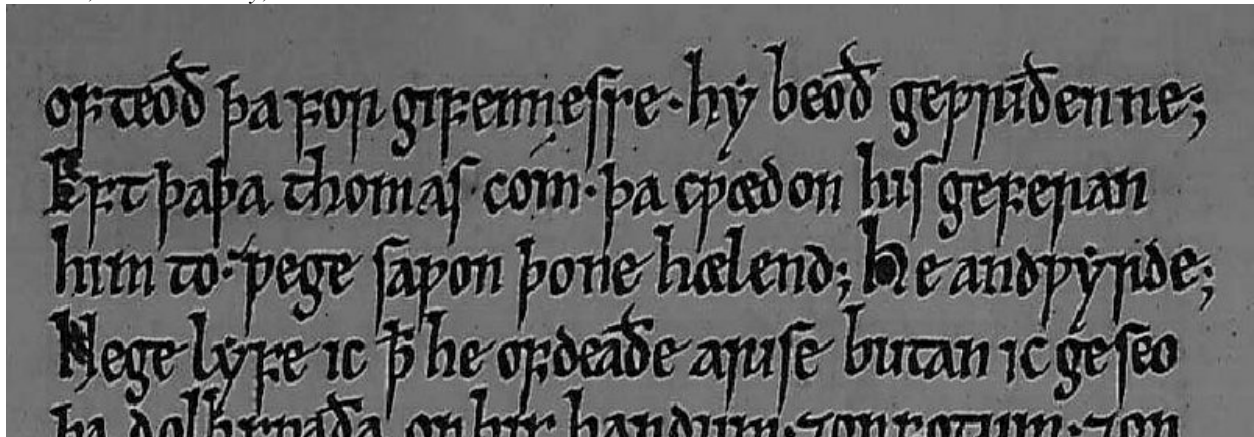
The texts correlate until after *forgifenysse* (spelled *forgifennesse* in the manuscript). At that point, the manuscript continues with "hy beoð gewriðenne" while the transcription continues with "þam bið oftogen." Clearly this is the beginning of a variant, but how long is the variant? Is there additional text in one or the other, or does the text branch here, continuing in parallel until matching up at some later point? This difference may be visualized in this figure (T=transcript, M=manuscript):

Additional Text in the Transcript:

```
<rdg wit="#DOE">********</rdg>
/
/
T: ********|************
M: --------|------------
```

Additional Text in the Manuscript:

```
T: ********|***********
M: ---------|-----------
|
|
<rdg wit="#N">-----------</rdg>
```

Parallel variants:

```
<rdg wit="#DOE">********</rdg>
/
/
T: ********|***********
M: --------|-----------
|
|
<rdg wit="#N">-----------</rdg>
```

This happens fairly often. The best way to figure out if the strange text is additional text or a parallel variant is to pick out some words or phrases from the strange manuscript text that are likely to be the same in the transcript as in the manuscript and search for them in the transcript (i.e. don't search for any words containing ð or þ; if a word has a "y" where you might also expect an "e" or "i" search for all variants).

In this instance, first I searched for "gewr" (the first part of "gewriðenne", the last word in the first line of the manuscript), and there was no hit in the transcription. So I went to the next line and searched for "gefer" (the first part of "geferan", the last word on the second line) and it came up in the same context in the transcript. From here I was able to track back through both texts and find the precise variation: one short parallel phrase ('þam bið oftogen;' in the transcription and 'hy beoð gewriðenne;' in the manuscript) followed by one longer phrase that appears in the transcription but not in the manuscript ('þa næs thomas þær an þæra twelfa apostola;'). Then it picks back up with only minor spelling variants ('ða þomas' vs. 'þa thomas').

Although you will see short variants (both additional text and parallel variants) it's highly unlikely that you'll find huge variants in the text (more than a few lines). It's more likely that you'll have sections that are rearranged; you'll have to do the same kinds of searches (look for words or parts of words that are in the same context in the trans. as in the ms. - it's a little like solving a puzzle) and mark them up accordingly.

The markup resulting from the variants above:

```
<app>
<rdg wit="#DOE">þam bið oftogen;</rdg>
<rdg wit="#N">hy beoð gewriðenne;</rdg>
</app>
</seg>
<seg n="230.4">
<note>John 20</note>
<app>
<rdg wit="#DOE">þa næs thomas þær an þæra twelfa apostola;<note>This
entire segment does not appear in the manuscript.</note></rdg>
</app>
</seg>
<seg n="230.4"><note>John 20</note><lb ed="#N" n="2"/>Eft þa
<app>
<rdg wit="#DOE">ða þomas</rdg>
<rdg wit="#N">þa thomas</rdg>
```

```
</app> com. þa cwædon his geferan <lb ed="#N" n="3"/>him to: we gesawon þone
hælend;
</seg>
```

1.3.5. and/ond vs. tyronian et (7)

In some of the manuscript we are editing, the scribes consistently have "7" (tyronian et) for "and," while the DOE text consistently spells this word out. If you find that the manuscript you are working with has "7" 99% of the time, do not take the time to mark the "7" symbols - this can be done later through a global change. This means that if you do come across an instance of the word "and" spelled out, you need to make a note of it! Use <note>your note here</note> here, or anywhere that you wish to note something curious, interesting or confusing in the manuscript.

1.4. Abbreviations

1.4.1. General

'*Identify all expanded abbreviations using the <expan> tag. Full descriptions will be added later*.'
<choice>, <expan>, and <abbr> tags. <expan> and <abbr> nest within <choice>. <abbr> contains the complete abbreviated word including specific indication of the abbreviated letter or letters, while <expan> contains the complete expanded word. For example:
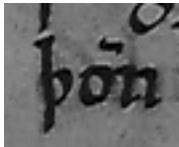
Would be encoded as

```
<choice>
<abbr>þ<hi n="abbr">o</hi>n</abbr>
<expan>þonne</expan>
</choice>
```

The complete abbreviated word 'þon' is in <abbr>. The "o", which has a macron over it in the manuscript, is noted using <hi n="abbr">. The @n attribute indicates that this letter is special ('highlighted') because it somehow represents an abbreviation. <hi> may be used elsewhere, with different values for @n, to note other letters or words that are special for some reason.

1.4.2. Attributes

on <expan>

<expan> takes an attribute @resp. For the initial encoding, where abbreviation expansions are taken directly from the transcription, the value of @resp should always be "#DOE":

```
<expan resp="#DOE">þæt</expan>
```

As the project continues and the initial encoding is expanded, the #DOE values here might be replaced or simply removed. See also the next section for those instances when the value of @resp may be other than "#DOE".

on <abbr>

<abbr> takes the attribute @corresp, which notes the description of the abbreviation in the <encodingDesc> section of the <teiHeader>. This links the specific instance of an abbreviation to a generalized description in the header. This saves us the time and effort of having to individually describe each individual abbreviation, ensures that abbreviation descriptions are the same for each homily and in each manuscript, and allows us to describe abbreviations in as much detail as required. It is also a simple matter to add new abbreviations or variants of existing abbreviations to the header as encoding progresses. This practice also opens up the option of linking the edited text to more detailed manuscript descriptions that could be developed and added to the project at a later date.

The contents of <abbr> should mimic as closely as possible the words as the appear in the manuscript, but using "plain letters" (i.e., "u" instead of "ū". For example, a crossed thorn should be represented simply by a thorn:

```
<abbr corresp="#cross-thorn"><hi n="abbr">þ</hi></abbr>
```

A "u" with a macron over it should be represented by an "u":

```
<abbr corresp="#macron-u">lazar<hi n="abbr">u</hi></abbr>
```

As the project progresses, we may choose to replace "plain letters" with letters that better reflect their manuscript appearence, though we should keep in mind that since this is an image-based edition, and the manuscript image will never be too far away, we need not concern ourselves with mimicking the manuscript completely (this would be a major distraction in my opinion!).

1.4.3. What Abbreviations?

Old English manuscripts contain very few abbreviations. The most common one by far is the crossed thorn (þ with a bar through the ascender). The @xml:id assigned to this abbreviation is #cross-thorn (for clarity, I've left out the <hi> element from these examples):
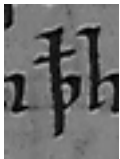

```
<abbr corresp="#cross-thorn">þ</abbr>
```

The other most common abbreviation is "þonne" (þon with a macron over the "O" and/or the "N"). The @xml:id assigned to this abbreviation is #þonne:
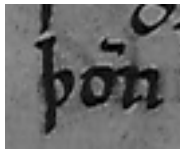

```
<abbr corresp="#þonne">þon</abbr>
```

You may also see macrons over various letter that represent a contracted "m" or "n", usually at the end of words. See the Abbreviation List in the header for a complete list of macrons.

If you run across something you expect is an abbreviation but it is not listed in the Abbreviation List, please contact me immediately: doporter@indiana.edu.

1.5. Abbreviations within Variants

You may come across instances where variants occur within abbreviated words. For example, the DOE might have "Þisum" whereas the MS has "ðisu" with a macron over the final u. In this case, you will need to record both the variant (between "Þ" and "ð") and the abbreviation and expansion of the macron-u. To do this, place the variant markup on the outside, and record the abbreviation and expansion within the manuscript reading. '*In this case, put your own @xml:id as the value of @resp for the expansion*.' If Doug were to record such an instance, his markup would look like this:

```
<app>
<rdg wit="#MS">
<choice>
<abbr corresp="#macron-U">ðis<hi n="abbr">u</hi></abbr>
<expan resp="#DR">ðisum</expan>
</choice>
</rdg>
<rdg wit="#DOE">
Þisum
</rdg>
</app>
```

1.6. Breaks in variants or abbreviations

At times, you will find page breaks or line breaks that come in the middle of words that are either variants or abbreviations (or both). Remember that breaks indicate what is happening with the text *on the manuscript*. So, when a page break occurs in a variant, place it only in the variant corresponding the the manuscript in question. Please do not add spaces before or after the line break:

```
<app>
<rdg wit=#DOE">helendes</rdg>
<rdg wit="#U">hæle<lb ed="#U" n="15"/>ndes</rdg>
</app>

<choice>
<expan resp="#DOE">þisum</expan>
<abbr corresp="#macron-U">þi<lb ed="#U" n="17"/>s<hi n="abbr">u</hi></abbr>
</choice>
```

1.7. Deletions

To note scribal deletions (either erasures, or crossouts, or underdots):
If you can make out the deleted text, use the <del> tag:

```
<del>on</del>
```

You may add a note indicating how exactly the letters were deleted, though this information will be added to the code later.

If you can't make out the deleted text but you can see that there has been an erasure or letters are scratched out, please leave a note.

1.8. Additions

Occasionally the scribe will add letters, words or phrases between lines or in the margins, or to replace deleted text. These additions should be marked with the TEI <add> tag. Place the added text where it makes most sense to you (whether in the middle of a line, the middle of a word, or at the end or the beginning of a line).
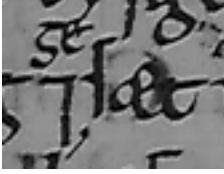

```
<add>ge</add>sæt
```

If you have additions that you are unsure about, just make a note, marked with your xml:id value from the header on the @resp attribute, that there is an addition in that area of the manuscript and another editor will check it for you:

```
<note resp="#DP">There is an addition in the margin here, but I'm not sure
what to do with it</note>
```

1.8.1. Marginal additions

Some of our manuscripts include marginalia added by later scribes (for example, the famed "tremulous hand" of Worcester). These marginal notes are also identified using the <add> tag, with the addition of the @resp tag to identify the noting scribe.

```
<add resp="#THW">marginal note</add>
```

If a marginal note appears in the left or right margin, append it to the beginning or end of the line to which the note refers. If at the beginning of the line, place it **after** the <lb/>
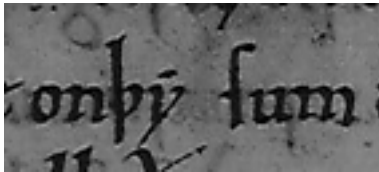
```
<lb/><add resp="#THW">marginal note</add>...<add resp="#THW">marginal
note</add>
```

If the note is a gloss, nest a <gloss> tag within the <add> tag.

```
<add resp="#THW"><gloss>marginal gloss</gloss></add>
```

1.9. Spacing

Ignore differences in spacing for now. If manuscript spacing is different from transcription spacing, use the spacing provided in the transcript.
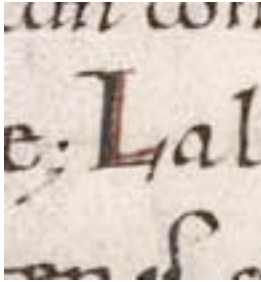

Transcription has:

```
on þisum
```

We have (including variant tags):

```
on <rdg wit="#DOE">þisum</rdg> <rdg wit="#U">þysum</rdg>
```
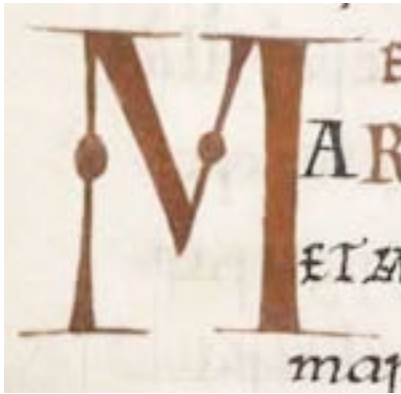
1.10. Special Characters

In some of the manuscripts, characters may be set apart by some kind of decorating or color filling. Please mark them as follows:

For filled letters:


```
<c rend="fill">L</c>
```

For colored letters:


```
<c rend="color: red">M</c>
```

(this format mirrors that of cascading stylesheets. Replace "red" with the color of the ink you are dealing with)

In cases of extreme decoration, accompany the text with a note:

```
<c rend="color: green">Oft</c><note>small "f" and "t" appear inside a large
"O"</note>
```

1.11. Damage

There are three main elements for recording damage: "damage", "unclear", and "agent".

On the one hand, use **damage** when text is damaged but still legible. Use the **agent** attribute to describe what has caused the damage on the page.

On the other hand, use **unclear reason="damage"** when text is completely illegible. As with damage, Use the **agent** attribute to describe what has caused the damage on the page.

It is to be assumed that the text inside an <unclear> tag **is not legible** in the manuscript and reflects accurately the text as found in the DOE edition. Text inside **damage** is from the manuscript as well as can be determined.

In the first example, the first two letters are damaged by a slit but the entire word, including "þy", marred by a slit, is still legible.

```
 <damage agent="slit">þy</damage>sum
```

In the second example, the first two letters are covered by adhesive and are completely illegible/lost. The letters in the tag are the same as those in the DOE.

```
<unclear reason="damage" agent="adhesive">þi</unclear>sum
```

Notes:

- **unclear** and **damage** may only have one agent per tag (and **unclear** may only have one reason per tag). If you have multiple agents, or reasons, you will need to nest multiple tags.

```
<unclear reason="damage" agent="frame"> <unclear reason="damage"
agent="MSedge"> w </unclear> </unclear>
```

- Values for **agent**: wear, adhesive, slit, smoke, MSedge, frame, hole, stain, tear (please add as you find them, and be careful to use only values from this list) **note that values for agent may not contain spaces**

1.12. Editorial Notes

As you are encoding, you will likely come across interesting, curious or confusing things in the manuscript, or things that you do not know how to encode. The <note> element is available to you to leave notes either for the technical editor (Dot) or the project director (Aaron). <note> can appear in most places, feel free to use it liberally.

1.13. Note Revisions

Supply a new change (in revisionDesc) every week. Do it more than that if you find it useful. You must do this at least once a week! Send me a version of your file every Friday afternoon so I can see how you are proceeding and check to make sure that your encoding is just fine.

1.14. Questions?

If you have any questions as tagging progresses, you can leave a note for me (Dot) in your XML file. Simply type out your question (adjacent to the area) and surround it with the note tag - see the example in Additions. Please also identify your note using the @rend tag, using the value of your xml:id (from the header) to show that this is your note. When I answer, my notes will be identified as @resp="#DP". If there is a particular word or phrase you're having trouble with, you can highlight it in the XML by surrounding it with the hi tag with attribute @rend value ed:

```
<hi rend="ed">weardmen</hi><note resp="#RU">what is this?</note>
```