CHAPTER 3:

How to Use

-- and NOT USE --

The University of Iowa TDM ETD Tags

This chapter provides the reader with the basic knowledge for using the University of Iowa Graduate College Document Type Definition (DTD) for Electronic Theses and Dissertations (ETD's), called Thesis and Dissertation Markup (TDM). Our TDM DTD uses a combination of academia-specific tags from the Text Encoding Initiative (TEI) and some HTML tags to enable the electronic thesis or dissertation to be viewed in older browsers. What we are doing is roughly describable as "XHTML" or, Extensible HTML.

These instructions assume you have a working knowledge of the basics of document markup technology (e.g., what an element is, attributes, open and close tags, empty tags, etc.). In addition, there are specific rules for XML which are more strict than the usual HTML markup. These strict rules are based on a notion of what is called well-formedness. Well-formedness is a new concept introduced by XML. Essentially this means that all elements must either have closing tags or be written in a special form (as described below), and that all the elements must nest. The following short review is derived from the World Wide Web Consortium's (W3C) web site http://www.w3.org/TR/xhtml1/.

- 1. Although overlapping is widely tolerated in existing browsers, it cannot be done in XML. This can also be called "straddling" your tags.
 - o CORRECT- nested elements:

here is an emphasized <emph>paragraph</emph>.

o INCORRECT- overlapping elements:

here is an emphasized <emph>paragraph.

- 2. Element and attribute names must be in lower case TDM documents must use lower case for all HTML element and attribute names. This difference is necessary because XML is case-sensitive e.g. and are different tags.
- 3. For non-empty elements (any element or tag which contains something -- see #5 below for explanation of an empty element), end tags are required. In HTML, certain elements were permitted to omit the end tag; with the elements that followed implying closure. This omission is not permitted in TDM. All elements other than those declared in the DTD as EMPTY must have an end tag. An example of an EMPTY tag is the line break tag which is simply the equivalent of a hard return to start a new line. It which is commonly seen in HTML code as

 hut must be written in TDM as

 hybrid to be written i
 - o CORRECT- terminated elements:

here is a paragraph.

o INCORRECT- unterminated elements:

here is a paragraph.here is another paragraph.

It is important to note that the latter incorrect code will "work" or view correctly in most browsers like Netscape, but will **not** pass the Graduate College Examiner's technical checks of your ETD. This means it will be <u>rejected</u> if you do this incorrectly.

4. Attribute values for elements/tags must always be quoted, that is, contained within single or double quote marks (whichever you choose, either "" or ", you must match them for each attribute, you cannot have "somthing', but instead, "something" or 'something'):

So, for working with attributes or, "variables" which are part of an element or tag, you should be careful to note this.

• CORRECT- quoted attribute values:

o INCORRECT- unquoted attribute values:

5. Empty elements must either have an end tag or the start tag must end with />. In TDM, we will only use the former. For instance, <hr /> **NOT** <hr>< /hr> (the tag for horizontal rules, those little lines or bars separating some text that look like this in a browser):

• CORRECT- terminated empty tags (note space before the "/"):

<hr />

o INCORRECT- unterminated empty tags:

 <hr>

As above with unterminated tags (#3), note that the latter incorrect code will "work" or view correctly in most browsers like Netscape, but will **not** pass the Graduate College Examiner's technical checks of your ETD. This means it will be <u>rejected</u> if you do this incorrectly.

- 6. Any time you have an attribute that is called "id" such as in the various div, div1, div2, div3, and div4 elements which structure the main and sub headings for your ETD, the value of that attribute must begin with a letter of the alphabet:
 - o CORRECT- an id beginning with an alphabetic character:

<div class="Chapter" id="chap3">

o INCORRECT- and "id" which doesn't begin with an alphabetic character:

<div class="Chapter" id="3">

This enables various sophisticated features of XML, and if you do not do this your document will not pass the technical testing by the Graduate Examiner.

7. (NOTE: unless you are using additional scripting and programming, you may ignore this point).

If you are using scripts such as Javascript, please note that the Graduate College and the University of Iowa cannot guarantee perpetual compatibility of these scripts and their functions with later browser technologies. It is up to your committee to determine if this potential--but not guaranteed--vulnerability will be a problem in later years. In TDM, the script and style elements are declared as having #PCDATA content. As a result, < and & will be treated as the start of markup, and entities such as < and & amp; will be recognized as entity references by the XML processor to < and & respectively. Wrapping the content of the script or style element within a CDATA marked section avoids the expansion of these entities.

```
<script>
  <![CDATA[
    ... unescaped script content such as JavaScript...
  ]]>
</script>
```

The nature of these tags and the basic structure for your ETD is explained in detail throughout this section. If you read and follow it carefully you'll produce an ETD which passes the Graduate College Thesis Examiner Standards, and which will stand the test of time and internet use.

How the Top and Bottom of the Document Must Look

The top and bottom of each file requires certain information. Different programs you use will do much of this for you, but there are peculiarities, so you need to be sure the final product has key information as presented below. You can check/change this by opening the files (and do this before final submission just to be sure) in BBEDIT (MacIntosh) or Wordpad--not Notepad--(Windows95-98).

The Top or Header Matter

Here is the basic information for inclusion in the header (you **must** use the "head, ""title," and "link" tags):

The top line will tell browsers about how the document is XML as well as how to read any special language characters. If you are working in a Non-Roman-Script language such as Japanese, use UTF-16 as the value for the "encoding" attribute in the <? xml ?> declaration. The second line with "DOCTYPE" identifies the basic root element (that element of which all others are a subset) which for UI ETD's is "html" (remember that XML is case-sensitive). Next comes the actual instance of that root element, "html". The "body" tag **must** have the "bgcolor" attribute with "#" and six (6) "f's" to be sure the background is white (otherwise browsers default to grey).

Then you need to identify whether it is the "front" matter, a chapter ("div") or "back" (with bibliography/"bibl", "notes" and "appendix"--if any). So, here are the possible lines of code which can come after the "body" tag (think of these as the grand backets around the very "outside" of each file you make):

- **<front>** (if it's the front matter file)
- <div class="Chapter" id="chap3"> (if it's a chapter file)
- **<back>** (if it's the back matter file)
 - **(if it's the bibliography file at the back this comes next)**
 - **(appendix)** (if it's the appendix file at the back this comes next, and so on).
 - **(if it's the notes file at the back this comes next)**

These are the different types of headers the various sub parts of the dissertation will have. In essence, there are three types of ETD files: the "front" file, the chapter/"div" files, and the back files with three sub-sections: bibliography/"bibl", "appendix," and "notes" (it is recommended that, for simplicity, all back matter be contained in one file such as "back.xml" unless you have extensive or large appendices or bibliography/notes). Each one, therefore, has different lines after the first four which deal with XML, DOCTYPE, the "html" root, and the "body" tag which are the same for all three types. A typical list of files, then, might be "front.xml," "div1.xml," "div2.xml," "div3.xml," "div5.xml" (each representing a distinct chapter), and "back.xml."

The Bottom Matter or, "Closing all the Top Tags"

As you know, XML requires all tags which "open" to also close (see above). Thus, in reverse order from what you put in the header matter, you close each tag at the very end of your file. So these are the five types of possible ending tags, depending on what file type you are writing:

- First you must close whatever of the five types you are working with:
 - </front> (if it's the front matter file this closes before "body")
 - </div> (if it's a chapter file this closes before "body")
 - </appendix> (if it's the appendix file at the back this closes before "back," and so on)
 </back> (then it's time to close the back matter file)
 - </notes> (if it's the notes file at the back this closes before "back")
 - \Box </back> (if it's the back matter file, this closes next)
 - </bibl> (if it's the bibliography file at the back or appendix or notes, that closes before "back") □ </back> (then it's time to close the back matter file)
- Then you close the "body" and "html" tags in reverse order from the header the same for all five types:

</body> </html>

Here is a "tree view" showing the general structure of a typical TDM ETD below. Be sure to note how the 1st, 2nd, and 3rd sublevels of the chapter nest within one another in descending order, and are parallel. Each one, however, you see has a unique value for the "id" attribute:

```
<title>
     Here is the Title of This Chapter
  </title>
  <link rel="stylesheet" type="text/css" href="tdm.css"/>
</head>
<body bgcolor="#ffffff">
        <div class="Chapter" id="chap1">
                  <div1 id="chap1.1">1st sublevel or section
                          <div2 id="chap1.1.1">2nd sublevel or section
                                  <div3 id="chap1.1.1.1">3rd sublevel or section
                                </div3>
                                  <div3 id="chap1.1.1.2">3rd sublevel or section
                                </div3>
                        </div2>
                          <div2 id="chap1.1.2">2nd sublevel or section
                        </div2>
                </div1>
                  <div1 id="chap1.2">1st sublevel or section
                          <div2 id="chap1.2.1">2nd sublevel or section
                        </div2>
                          <div2 id="chap1.2.1">2nd sublevel or section
                        </div2>
                </div1>
        </div>
</body>
</html>
```

Note: You will have noticed that the "XML" and DOCTYPE parts of the top header matter <u>do not</u> have close tags. You can simply cut and paste these from the demos above if you want to be sure you have it right.

General Overview of What Tags are to be Used and Where According to the Three Major body tags, "front," "div," and "back"

In this section we will look at general rules for what tags and kinds of content go where. Some tags only allow typing of keyboard characters between them (technically called "#PCDATA" or "Parsed Character Data"). Others allow specific tags while not allowing others (this is called a "content model" and most XML software obeys these rules or, at least, displays them). It is <u>very</u> important that you pay close attention to what tags are allowed where, because this is one of the main things checked when you submit to the Graduate College, and will result in rejection of your ETD if not done correctly.

"Tags" are, in essence, the same thing as "elements" inasmuch as a tag defines a specific element of the ETD, whether a "titlePage" or a paragraph/"p" or something in bold/"b", etc. The root element, or tag at the "base" of each of your files is the "html" tag. It is the "parent" of all other tags because all other tags are contained in it. Thus, to explain the relationship of a tag to those tags contained in it, or the relation of a tag to the tag which contains it, we call them child and parent tags. A parent tag contains other "children". Some parent tags have very specific tags belonging to them, as with the "front" tag which has specific children found only in the front matter (such as your dissertation or thesis certificate of approval, abstract, table of contents, etc.).

The main limitations are discussed below according to the subsections they below, for <u>front</u>, <u>chapters/"div"</u>, and the back matter with the three kinds of back tags, <u>bibliography/"bibl"</u>, <u>appendix</u>, and <u>notes</u>.

Tags which are Used for the "front" Matter

The front matter, more than any other area, has the most specific and inflexible set of tags. Most software you use will coach you through this process, and there is also a demo file of front tags to get you started (see printout at the end of these instructions if you are reading them in hardcopy).

IMPORTANT: To make the table of contents links work, you need to <u>create an anchor</u> with the "a" tag and the name attribute for each section and subsection you wish to have accessible from the table of contents. <u>HINT</u>: Any time you make a new "div1" or "div2" or "div3", make an anchor just before it if you want the table of contents to reference it. Here's how it was done for the section of these instructions you are reading now:

</div2>

You can either use the file you got from the Graduate College and cut and paste these sections into a new "front" file of your own (a pointless reinvention of the wheel unless you **really** like doing everything yourself), or just edit the file you received by adding your own personal information (replace the text the department, date, name, title, etc., with your own). You can also copy the code and sections at the end of these instructions if you are reading it in print form.

The following table shows the tags which are found <u>only</u> in the front matter. These are displayed according to which tags belong to which "parent" tags, with "front" being parent of all. There **are more** tags than these, but these are those which **should not** be used elsewhere than in the "front" file matter. Those with a question mark "?" are optional. The tags are shown in order from left to right ("front" coming first) and top to bottom ("titlePage" is the first tag/element under "front" [front is the "parent" of titlePage] and it contains "docTitle," "docAuthor" [that's you!], "docDate," and "thesisadvisor" as its children).

	titlePage	docTitle		
		docAuthor		
		docDate		
		thesisadvisor		
	thesiscopyright?			
	certifapproval			
	dedication?			
	epigraph?			
front	acknowledgements?			
	abstract	abstractcover		
		abstractext		
		abstractapproval		
		(includes "thesisadvisor")		
	toc			
	table of contents			
	tableist			
	figurelist			
	symbolabbrevlist			
	preface			

This is a sort of "tree" showing how the front-only tags are arranged according to which "parent" elements they belong to. Thus, "titlePage," "thesiscopyright," "certifapproval," "dedication," and the other elements in that verticle column are all children of the parent "front," and are "nested" within the front. Some of those, like "titlePage" have their own children (they are, then, sort of "grandchildren" to the "front" tag, but that is not their technical term). Obviously, of course, the "abstractext" element will contain sub-elements ("great-grandchildren" of the front, if you will) such as paragraph/"p" elements but these do not need to be listed here as the software will coach you on what to put where, and the demo file and code can assist for reference.

Tags which are Used for the Main Body of the Text: The "div"/Chapter Section Tags

The chapters, following Text Encoding Initiative (TEI) standards, are designated with the "div" tag, which has preset "class" attribute of "Chapter", to which you add the number as the "id" attribute beginning with one or more letters of the alphabet. To review what was discussed above, note the following example:

```
<div class="Chapter" id="chap5">
```

[then all your chapter five wisdom goes here]

</div>

Following the closing "</div>" tag, you would then close the remaining page tags as described above in "<u>Closing</u> <u>all the Top Tags</u>."

The main focus in this section, however, is on what goes <u>between</u> the <div></div> tags. There are three main subtags, each denoting a respective sub-heading level (each "id" must be a unique value, in other words, there can not be two "<div2 id="chap1.1.1">" tags, but there can be two --or many-- "div2" tags with different id's):

```
<divl class="subheading" id="5.1">
5.1
[for instance, Chapter 5, subsection 1, your wisdom goes here]
<div2 class="subsubheading" id="5.1.1"></div2>
5.1.1
[then, a sub-category of that wisdom goes here]
<div3 class="subsubsubheadhing" id="5.1.1.1">
5.1.1.1
[and an even further sub-divide of your wisdom
would go here]
</div2>
</div2>
</div1>
```

The nice thing is, the contents of each is fairly unrestricted, and quite simple to use for a variety of formats and needs.

IMPORTANT: (repeated from above for emphasis) To make the table of contents links work, you need to <u>create an anchor</u> with the "a" tag and the name attribute for each section and subsection you wish to have accessible from the table of contents. <u>HINT</u>: Any time you make a new "div1" or "div2" or "div3", make an anchor just before it if you want the table of contents to reference it. Here's how it was done for the section of these instructions you are reading now:

```
<div2 class="Chapter_div_tags" id="3.2.2">
       <a name="divtags" class="anchor">
               <center>
                       <font size="4">
                                 Tags which are Used for the Main Body of the Text:
                                                <br /> The "div"/Chapter Section Tags
                       </font>
               </center>
       <br />
       The chapters, following Text Encoding Initiative (TEI)
                     standards, are designated with the "div" tag, which has
                     preset "class" attribute of "Chapter", to which you . .
       (. . . the text then continues until the conclusion
               of the section, with several paragraphs and so forth contained in it.)
```

```
</div2>
```

You can either use the file you got from the Graduate College and cut and paste these sections into a new "front" file of your own (a pointless reinvention of the wheel unless you **really** like doing everything yourself), or just edit the file you received by adding your own personal information (replace the text the department, date, name, title, etc., with your own). You can also copy the code and sections at the end of these instructions if you are reading it in print form.

Tags for the Back Matter

The back matter can have one of three kinds of content, and accordingly the tags will have one of three names:

"bibl"/bibliography, appendix (if any) and "notes." If you have long Appendixes or Bibliography/notes, you may want to divide the back matter into two or more files, such as "bibl.xml," "notes.xml," and so forth. If you're putting all the back matter in one file, the main tag codes would look like this:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html SYSTEM "tdm.dtd">
<html>
 <head>
  <title>
       Back matter for my ETD
  </title>
  <link rel="stylesheet" type="text/css" href="tdm.css"/>
  </head>
        <body bgcolor="#ffffff">
                <back>
                        <notes>
                                         Your numbered notes list would go here
                         </notes>
                         <bibl>
                                Your bibliography would go here
                         </bibl>
                        <appendix>
                                         Your appendix (or appendixes) would go here
                         </appendix>
                </back>
        </body>
</html>
```

```
The Bibliography/"bibl" back Matter Tag
```

The bibliography can be done one of two ways, with either the definition list/"dl" format, or the unordered list/"ul" format. These tags are discussed with more detail <u>below</u>. In this section, we will look at the basic manner in which the "bibl" section of the back matter is to be formatted.

Here's the code if you're making the bibliography a separate file. Prepare the top tags as described <u>above</u> for the XML and DOCTYPE declarations, plus the "header" and "title" tags to identify the file (the "title" and "head" are optional, but recommended):

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html SYSTEM "tdm.dtd">
<html>
<head>
<title>
BIBLIOGRAPHY
</title>
<link rel="stylesheet" type="text/css" href="tdm.css"/>
</head>
<body bgcolor="#ffffff">
<body bgcolor="#fffffff">
<body bgcolor="#ffffff">
<body bgcolor="#ffffff">
<body bgcolor="#ffffff">
<body bgcolor="#ffffff">
<body bgcolor="#fffff">
<body bgcolor="#ffffff">
<body bgcolor="#fffffff">
<body bgcolor="#ffffff">
<body bgco
```

Then your bibliographic entries are made with either the definition list/"dl" format (see section <u>below</u> on "Tags Containing Only Typing"):

<dl>

<dt> Gardner, John Robert. </dt>

<dd> The Developing Terminology for the Self in Vedic India . Ann Arbor: University Mircrofilms, International, 1998. </dd>

</dl>

You will notice that the "author" tags are not allowed in this format so, if you wish to use these tags, the unordered list/"ul" should be used instead:

Notice the "worktitle" tag. **ALWAYS** use this to cite works' titles, NOT the title tag. Since HTML limits "title" to the "head" element, for now, all titles you wish to identify other than "docTitle" (the title of your dissertation or thesis), will use "worktitle." The "level" attribute signifies whether it is "m" (monographic- a book, monograph, or other publication under a single autonomous title), "s" (a series title), "j" (a journal title), "u" (unpublished title, like a dissertation ;-). Additionally, you can use "class" to specify it is an "abbreviated" version of the title, if it is the "main" title, if it is a "subordinate" title, or, if it is a translation, for instance, it is a "parallel" title. You can choose the format that looks best for you. You can also refer to the section <u>below</u> describing the list tags in more detail. Always remember to close the tags at the end of the file which you began at the top (see above) if you're making separate files for the back matter:

</bibl> </back> </body> </html>

In the bibliography/"bibl" section, whether you use the definition list/"dl" or the unordered list/"ol", the individual elements of the bibliography entries should be tagged with specific identifiers for both searching and also for later formatting for print if desired (consult with your dissertation/thesis committee to be sure). You actually are not *required* to use all of these. Here is the entry above, showing the use of TEI tags to distinguish components.

<author>Gardner, John Robert</author>. <worktitle>The Developing Terminology for the Self in Vedic India </worktitle>. <pubPlace>Ann Arbor</pubPlace>: <publisher>University Microfilms, International</publisher>, <date>1998</date>.

There is also an "editor" tag for these entries which is also a TEI tag. For articles where a citation includes a range of pages, there is not a TEI tag, but an optional "pages" tag is allowed if you need to be this specific.

The Appendix/"appendix" back Matter Tag

Another kind of back material is the appendix. This is, of course, optional because you can have one, several, or none. Depending on their length, they can be in single files one per appendix (e.g., "appendix1.xml", "appendix2.xml") or all in one (e.g., "appendix.xml"). Here's how you'd do it if the appendix were a separate file. The top tagging is the same as the other back matter (see <u>above</u>) with the exception of the "appendix" tag in place of the "bibl" tag (use the "id" attribute to indicate which appendix if there is more the one):

The appendix can accept all the regular tags valid for the chapter/"div" sections, so refer to the discussion <u>above</u> for the general summary and look <u>below</u> for specifics on the different tags. Always remember to close the tags at the end of the file which you began at the top (<u>see above</u>): </appendix </body> </html>

The Text Encoding Initiative/TEI tag for tracking an argument, the "argument" tag is used **only** in the appendix. If you are tracking an argument in the chapters/"div" section with "hi" you must switch to "argument" when writing in the appendix. See <u>below</u> to learn more about the TEI tags.

The Notes/"notes" back Matter Tag

This is the file where your notes are placed. It is easiest to not rely on the note numbers for linking to them, the anchors can be named, e.g., "about_flies" and "Adams_article_94"-- therefore during drafts, if your committee wants traditionally-numbered notes, you don't have to change anchors and links every time the order of the text and the cited notes changes, all you have to do is change the numbering. The numbers are for your reviewers to refer to them in comments. In the text, citing notes is best achieved by links. This is an electronic rather than a paper medium. The original printed footnote/endnote "look" was to facillitate reference to ancilliary or additional related materials. In an ETD or any other meta-tagged electronic text, links do this. The convention of superscript numbers is superfluous to the function of the links in the electronic text medium. Thus a "footnote" here would be like this [note14]. This is a convention currently in widespread use on the net.

Also, it is very important that your set the "class" attribute for the "a" for the notes correctly. You remember that you must identify an "a href" that goes to the notes with the "class" set as "note", but in the "notes.xml" file itself--**and ONLY in the notes.xml file**--you must also use "note" as the "rel" for the anchors you make with the "a" tag as the destination points of your notes: and so on for each note. Remember, though, this is only in the notes.xml file that the "a name" tag has "note" as the "class" otherwise, of course, "a name" has "anchor" for its "class."

As with the other back sections, the top of the file must be specifically identified:

<notes>

Then the ordered list/"ol" works easiest to arrange the notes. To get a space between each note, use the "br" tag twice as shown immediately below:

- 1. Here is text for a note.
- 2. Here is text for another note.

Here's the code for the notes list:

<01>

```
Here is text for a note. <br /><br />
Here is text for another note.<br /><br />
```

You can also refer to the section <u>below</u> on the list tags in more detail. Always remember to close the tags at the end of the file which you began at the top if you're making it a separate file:

</notes> </back> </body> </html>

Tags which can **only** contain typing (no other tags)

Some tags can <u>only</u> contain typing. The columns below separate what tags can be used in which sections (though not <u>all</u> tags that might be used in each of the three major dissertation subdivisions are listed, only those which are "text/typing-only") that permit only typing. No other tags can be placed in the following tags:

Tags Containing Type Only According to Dissertation Section					
front	div	back			
docDate	caption (in a table)	pages			
docAuthor	author	publisher			
thesisadvisor	dt	editor			
sup, sub	sup, sub	pubPlace			
author	pages	volumeissue			
***	date	date			
***	th	dt			

To find out what other tags can be used in each section, <u>refer to the section above</u> on "Major Tags and Their Contents."

Detailed Explanation of Key Tags

In this section the four most complex tags--in terms of their sub-tags/elements or children as well as their attributes--are discussed. These tags are those for making links/"a", for inserting images/"img", for making tables/"table", and for the list tag. In addition, there is a summary of the remaining tags which are not quite so complex, but are nonetheless very common.

Using the "a" Tag

You have probably noticed by now (or have known for a while) that some tags have other information besides the tagname in them. For instance, if you were to make a link going to the <u>UI web page</u>, you would see that this section has an "a" tag and something called "href" and "class" after it:

UI web page

"href" and "class" are called attributes, and they give the tag specific characteristics and specify any instructions you want carried out by the browser when someone reads your ETD.

For instance, to set an anchor within a document: , in order for the link to lead not just to a whole file but to a particular place in the file (such as a particular subsection of your Chapter 3 --for instance, the "div3.xml"-- file), you use the "a" tag and the "name" attribute. An easy way to remember this is, "if you want to go to a specific place in a file, you have to give it 'a name" -- get it? Files have to have names in order to be found on the WWW by browsers, and places within those files, then, have to have names too. You **do not** use the "href" attribute when making an anchor, you use the "name" attribute instead, with the "class" attribute set as "anchor". **Remember:** if you do not add the "class" attribute, clarifying how it is you are using the "a" tag, your files will fail The Graduate College Thesis Examiner's tests.

We use the traditional HTML syntax for anchors and links, designated by the attribute values "href"/for a link or "name"/for an anchor. In order to be sure and distinguish between internal cross references there is an additional required attribute which explains how you are using the "a" tag, and it is called the "class" attribute which has the following possible kinds:

- "ref"- when the "href" attribute is used to point to an anchor or file elsewhere in your own ETD
 - Such as a link to another point in a file: tables. NOTE: it is important that when you use an anchor in an "href" attribute value, it <u>must</u> begin with the ''#'' mark.
 - Such also, then is the case with a link to another point in a different file of the ETD, such as a link to a section of some author like Jan Gonda's works in the bibliography/"bibl.xml" file: Gonda link
 - Also, if you have multi-media files (MPEG's/".mpg"), you use the "a" tag to call for them: a sound file, with the "class" attribute set to "ref."
- "link"- when "href" is used to point elsewhere on the web (generally, when you use "http://www.something.org" as the value for "href"
 - Such as a link to the World Wide Web Consortium site. Note that the "class" attribute says "link."
- "anchor"- when "name" is used to set a place within your files where you want to link to. No material should go between these tags.
 - Such as the "anchor" established by
 - **NOTE**: When you are making an anchor in the "notes" file of your back materials, you should set the "class" attribute as "anchor." Note also that an "a" tag when used as an anchor doesn't contain anything--no typing or other tags--but MUST have a closing tag all the same.

- "note", uses "href" to point to an anchor which designates one note among all the notes in the back matter: [note23]. When setting the destination for
 - You can do a superscript, but the code is more complicated. Such as, if this were a section of your ETD, you might have a footnote with discusses how to get additional information about how to use the conversion program for word processor files to make them into XML, called Majix:

¹.

This "class" or relation attribute is required in all cases. This enables later XML interpretation of the kind of "a" tag that is, as well as vital formatting.

One other attribute is allowed which should be used sparingly. If a link goes someplace which you want to have open at the same time as the window in which you clicked ot activate the link (in other words, open two browser windows at once) you can use the "target" attribute--with only the value "new"--to open a new browser window. This link to the demo file does this. Here's what the code looks like:

 This link to the demo file does this

(but this one is for code demo only, it doesn't "work")

Remember, your reader can get confused if you use "target" to "spawn" too many windows. It's also an unspoken "no-no" on the 'net if you do. In summary, then, the "a" tag does several things (all of which require use of the "class" attribute):

1. It can make a connection to someplace else in your ETD with the "href" attribute:

some linked word or image goes here

NOTE: this example is a link not only to the particular file called "my_etd_file.xml" but also to the section within that file designated by an anchor called "section" (cf. #4 on anchors).

2. It can make a connection to someplace outside your ETD with the "href" attribute:

some linked word or image goes here

3. It can make a connection to one of your notes--whether you number them or otherwise--with the "href" attribute:

some linked word or image goes here

4. It can create and "anchor" so that a link can be made to that specific spot from someplace else-- note, nothing goes between these tags:

```
<a name="somename" class="anchor"></a>.
```

5. When making an anchor in the notes file the "class" attribute also reads "anchor":

```
<a name="note3" class="anchor"></a>
```

- 6. This is also how to set a link to a sound or movie (MPEG/".mpg") file.
- 7. With the "href" attribute, the "rel" attribute (always required) and the "target" attribute (optional, use with caution) you can open a link in a new browser window.

Also, ANY material can go inside the "a" tag, except when it is used as an anchor (with the "name" attribute), it is

ALWAYS empty (nothing between the open and close):

Images can even be put inside the "a" tag. It is also with the "a" tag that you can make an e-mail window pop up:

Email Me!

Note that when used for mailing, the "class" attribute should read "link" because, necessarily, the e-mail sent will go "outside" the ETD to elsewhere on the web.

Using the Image/"img" Tag

This tag is used to insert graphic images (JPEG's). It has two required attributes, the "src"/source (what the name of the image file is), and "alt" (where you type in a description of the image for visually impaired individuals and browsers without graphic display).

This tag basically places an image in the document. The location of the image, in your thesis/dissertation directory, is the required "src" element (e.g. 'source="picture.jpg"). in order to be compatible with visually impaired readers, or render some meaning in non-graphic renditions, the "alt" is also required (e.g. 'alt="picture showing peptide reaction of control group"). Width and height adjustments are allowed, but not recommended. Real sizing of images should be done before final placement as this enables them to be as memory-small as possible. Width and height do not change the file size of the image, only its display size, and they can distort it too. To change the actual **file** size, you need to work in an image preparation program like Photoshop.

"Border" is the band around the image. If it is set to zero, there is no border. "Align" can be top (the top of the image 'hangs' from the current line of text), bottom (the bottom of the image 'sits' on the current line) left- a useful option for layout control (moves the picture to the far left and any text will then use the right edge of the image as the left margin not unlike a text wrap), right-(mirrors what "left" does, only lines up on the right), and middle (the middle of the image aligns or 'straddles' the current text line). "hspace" and "vspace" are horizontal and vertical invisible 'buffers' or 'blank space/padding' around the image, given in number of pixels (10 is about 1/4 to 1/3 of an inch).

The "img" tag is what's called an EMPTY tag-- it does not have a closing tag, there is a closing mark contained in it. **IMPORTANT**The style sheet for ETD's which gives double-spaced text <u>requires</u> that all images be placed **between** paragraphs, or they will not display properly. Be sure to end the paragraph you are in with a tag before inserting the "img" tag. Here is an example of the tag with only the required codes:

Here is a detailed summary of the "img" attributes, taken from the <u>W3 international standard</u> (note, the "new" target attribute is used for that link) for the World Wide Web:

src

This attribute is required for every IMG element. It specifies a URL for the image resource, for instance a GIF, JPEG or PNG image file.

alt

This is used to provide a text description of the image and is vital for interoperability with speech-based and text only user agents.

align

This specifies how the image is positioned relative to the current textline in which it occurs:

align=top

positions the top of the image with the top of the current text line. User agents vary in how they

interpret this. Some only take into account what has occurred on the text line prior to the IMG element and ignore what happens after it.

align=middle

aligns the middle of the image with the baseline for the current textline.

align=bottom

is the default and aligns the bottom of the image with the baseline.

align=left

floats the image to the current left margin, temporarily changing this margin, so that subsequent text is flowed along the image's righthand side. The rendering depends on whether there is any left aligned text or images that appear earlier than the current image in the markup. Such text (but not images) generally forces left aligned images to wrap to a new line, with the subsequent text continuing on the former line.

align=right

floats the image to the current right margin, temporarily changing this margin, so that subsequent text is flowed along the image's lefthand side. The rendering depends on whether there is any right aligned text or images that appear earlier than the current image in the markup. Such text (but not images) generally forces right aligned images to wrap to a new line, with the subsequent text continuing on the former line. **Note**: some browsers introduce spurious spacing with multiple left or right aligned images. As a result authors can't depend on this being the same for browsers from different vendors.

width

Specifies the intended width of the image in pixels. When given together with the height, this allows user agents to reserve screen space for the image before the image data has arrived over the network.

height

Specifies the intended height of the image in pixels. When given together with the width, this allows user agents to reserve screen space for the image before the image data has arrived over the network.

border

When the IMG element appears as part of a hypertext link, the user agent will generally indicate this by drawing a colored border (typically blue) around the image. This attribute can be used to set the width of this border in pixels. Use border="0" to suppress the border altogether. User agents are recommended to provide additional cues that the image is clickable, e.g. by changing the mouse pointer.

hspace

This can be used to provide white space to the immediate left and right of the image. The HSPACE attribute sets the width of this white space in pixels. By default HSPACE is a small non-zero number.

vspace

This can be used to provide white space above and below the image. The VSPACE attribute sets the height of this white space in pixels. By default VSPACE is a small non-zero number.

usemap

This can be used to give a URL fragment identifier for a client-side image map. element.

The primary attributes you will use, however, will be the "src"--which is required to identify which image file you are inserting--and the "alt"--which is required to give verbal description of the image in case it cannot be seen for some reason. Use these two to get started, then add the others only as needed in your final drafts. This will eliminate errors.

Using the List Tags

These are very handy tags as they allow you to make nice pre-formatted lists, with indents and spacing, automatically with little trouble. There are three kinds, the ordered list (numbered)/"ol", the unordered list/"ul", and the definition list/"dl". Both the ordered and unordered lists ("ol" and "ul") contain the "list item/li" tag which, in turn, contains any possible content you wish to place there-- images, links, etc. The definition list contains the term defined/"dt" (a typing-only element, see <u>above</u>), and the definition itself/"dd." A summary of lists is included, with examples, below.

The Unordered List

Unordered lists are very common, they are, in fact, created in the order that you write, but they have only bullets to mark each entry, not numbers or letters as with ordered lists. They are very flexible. For instance, you can embed lists within each other. The attribute for bullet shape is "type" for either dots/"disc", squares, or circles for the bullet markers and "compact" to compress the spacing.

- □ Here's a list item/"li" in an unordered list
- □ Here's another list item/"li" in an unordered list, which contains an embedded list:
 - - This is a list item of an embedded unordered list
 - This is another list item of an embedded unordered list. Note that I have to close this current list item/"li", the embedded unordered list/"ul", and the list item/"li" in which it is embedded:
- And here's another list item at the main list level

That is a simple unordered list. It works well for the bibliography and most uses. Be sure to close tags properly when embedding, it can get confusing.

The Ordered List

Here's an ordered list/"ol". They can be nested too, with letters for sub-list elements, etc. You can set these by adjusting the "type" attributes. The automatic setting is regular numbers if you choose not to use the "type" attribute.

- a. <Ii> This is an ordered list, with the type attribute set to lower case letters instead of numbers. </Ii>
- b. You can embed these lists too.

- i. ii iii ii ii
- ii. <**li**> I can also sub-sub-embed a list. Note that the code doesn't specify "type," and I automatically get numbers.

<0l>

- 1. 1. Ii> This is a list item in the subsublist, no type attribute set
- 2. This is another list item in the subsublist, no type attribute set. Note that I will have to close this list item, the embedded sub-subordered list, and the list item in which it is embedded as well.
- iii. Here's an item at the sublist level. Note that I will have to close this list item, the embedded subordered list, and the list item in which it is embedded as well
 - </11>

- c. <**li**> And here's a final item at the main list level. It is closed, and the original ordered list is closed thereafter.

Ordered lists have several attributes you can adjust. Number sequences are automated, in order of each "li"/list item entered. But the "start" atribute lets you set the starting number or letter. The "type" attribute sets numbering

style: 1/Arabic numbers (the default setting, a/lowercase alphabet, A/uppercase alphabet, i/lowercase roman, I/uppercase roman.

The Definition List

Here's a definition list:

<dl>

<dt> Term to define </dt>

<dd> And here's the definition of it. Well, not really, but this is where the definition of the term would go. </dd>

<dt> Another term </dt>

<dd> And here's that other term's definition. Again, not really, but I had to find something to write for the example. </dd>

</dl>

Note: be sure to watch your open and close tags when doing embedded lists.

Using the Table Tags

This is perhaps the most complex--depending on how much organization you're seeking to impose on your data, images, or media (all of which can be put in tables under the "td" element)--and powerful tool. It allows very precise arrangement of text, graphics, and visual media. Table sections will size themselves to fit what you put in them, but you have some controls over this--you can't make a table segment smaller than its contents, of course. If you set your "border" attribute to "0", there is no visible table on your screen, just nicely arranged materials.

As noted, the "border," when set at zero, is invisible (see <u>Table 5 below</u>). At a value of "1", it is a small bar and anything above a value of 1 is a sort of three-dimensional shadowed relief. The "cellpadding" and "cellspace" variables determine how much space is allowed between the contents of each table cell and its border or how much space there is between the cells, respectively. Inclusion of the "caption" element is especially convenient for captioning images when they are placed with the "table" and "td" elements. You can also determine how much, in other words, what percentage of the browser window is taken up by using the width attribute with a "%" value (e.g., a width of 50% will make the entire table as wide as half the width of the browser window). NOTE: different users and different computers open their browsers different amounts, so this is a relative, rather than an absolute proportion.

There are three main sub-elements of a table, the caption, the table row/"tr" and the table cell/"td." The caption can be aligned at the bottom or top (i.e., centered below or above the entire table) of a table, and is thus useful for controlling precise layout of images and their captions. The table cell and the table header/"th" elements are also sub-elements of the table row/"tr." The table row/"tr" can include most kinds of content, but must have an individual table cell/"td." The "align" attribute centers all the content of the given row wherein it is invoked from top to bottom. There is the <u>typing-only</u> element called the table header/"th" which makes bold sections of text in the table.

The table cell/"td" is the most versatile and adjustable element of the table. It can contain just about anything, and accepts attributes governing its content and relation to other cells. It accepts "align" attributes. It can be used with the attributes of "colspan" to take more than one column (make itself wider), and/or "rowspan" to take up more than one row (make itself taller). The "tr" is the basic element forming the structure of almost all tables.

Here is a simple table with a border of 1 (to remove the borders, or "frame", set the "border" attribute to zero, per the code immediately below):

Some text can go here. Maybe a picture could go here

Some text or a picture or pictures can go here.

Now here is the code for that table:

```
> (this is for the first or top row)
            (this is for what goes in left-hand side of the top r
                <b>Some text can go here.</b>
            (this is for the right-hand part of the top row)
                <b>Maybe a picture ("img") could go here</b>
            (this ends the top row)
     > (this starts the second row)
            (this makes the single cell of the bottc
                 makes it go all the way across the table)
                <b> Some text or a picture or pictures can go here.
```

Remember to adjust cellpadding, cellspacing, and border according to what makes the best layout for your purposes. Copy the example above to practice.

However, there is no better summary than that at the <u>netscape site</u>, I've been using it for over 4 years now, and it covers everything I've needed, including the tables in this document. There is also a great summary, <u>at the same site</u>, of just the attributes and what they do. Print it out, you'll be glad you did.

Using the Other Tags

There are quite a few tags remaining, though many of these are text-only tags, such as those <u>described above</u>, allowing only typing. The remaining tags fall into two broad groups, those for formatting (coming from HTML), and those for marking content (coming from the Text Encoding Initiative/TEI). The TEI tags are **extremely** important as they give depth, dimension, and multiple avenues of access to your work for researchers. The HTML tags largely allow specific appearance adjustments for text such as bold, color, etc.

Other HTML Tags

There are quite a few HTML tags which affect appearance. The simplest way to summarize them, is to show them below with the way in which they affect the text. The table below includes the tags which shape your text.

bold	<i>iialic </i>	<u><u>underline</u></u>		
There is a ^{superscript} tag There is also a _{subscript} tag				
 Font for Size and Color	Font has a SIZE attribute, with values of 1-6			
	Also font has a color attribute, this one is navy			

Table 4: Other HTML Tags for Text

The following table offers values for the main colors you might wish to use for your text. Remember, some of these do not have a lot of contrast for white backgrounds and, so, should be used only with caution as folks with black and white screens couldn't read easily. Remember, each "#xxxxxx" value below is a value to go inside the quote marks of the attribute "color" in the font tag. You **must** have 6 digits or the whole tag won't work. Also, you must make sure that each value begins with the "#" mark, then the six digit/letter color value.

IMPORTANT (and good news!): if these codes are a bit tedious, you can just use the word that you see next to each one as the attribute value, without the "#" if this range of colors meets your needs. If you want subtle variations, you'll have to play with the six-letter values for each of these.

black = #000000	green = #008000	<pre>silver = #c0c0c0 (silver)</pre>	lime = #00ff00
gray = #808080	olive = #808000	white = #ffffff but you couldn't read it!	
maroon = #800000	navy = #000080	red = #ff0000	blue = #0000ff
purple = #800080	teal = #008080	fuchsia = #ff00ff	aqua = #00ffff

Table 5: The Color Attribute Values for Font

You will notice that no number values above are outside the 0-8 range, and no letters are outside the a-f range. Be sure that you do not asign values outside that range as your document will not work properly--and may not even display any text with a font color value outside those ranges.

There are also a few layout tags. The most common, of course, is the one you'll use over and over, the paragraph tag, which is also a TEI/Text Encoding Initiative tag:

Vrite your wisdom, include images and media, add tables or whatever, in here

Also, on those occasions where you are citing a section of text from another source, more than just a few words, there is the blockquote element. The blockquote element can also have a quote element within it, the "q" tag so that you mark the segment with a TEI tag as well as the HTML/block formatting tag (see below re. TEI tags for

"q".

dockquote>Here is a block of text. It's not really a quote from another text, it's just some stuff
I wrote so that you could see what the "blockquote" tags do to the layout of your document
when you use them to offset a quote from another source. </blockquote>

There is also a tag for setting exactly every space and line return/carriage return/hard return exactly as you type it, this is the "pre" tag for 'preformatted.

When you preformat, every
space and hard return
that you type is preserved.
but only in monospaced courier text.

The ETD style sheet will render double-spacing of paragraphs.

The only remaining HTML layout tags are for the general layout around and inside the paragraphs. The most common is the "center" tag:

<center>this is centered</center>

You can also make a "horizontal rule" to put a bar to separate sections, it can be adjusted for width and thickness with the width and size attributes.

this horizontal rule above set the width at 200 pixels with this code:

<hr width="200"></hr>

this one above set width and size with this code:

<hr width="150" size="12"></hr>

Finally, there is the line break, or hard-return tag, to make a line double space (don't forget the space before the "/" as discussed above:

</br>).

The Text Encoding Initiative (TEI) Tags

These tags come from the Text Encoding Initiative's academic tag sets, established as long ago as 1989 (which is a long time ago in computer development terms! -- no one even imagined the world wide web back then). Many of the tags in the front matter are TEI tags, and many of the tags in the back matter (the bibliography/"bibl" section) are TEI. Here is a table of the main subdivisions of the ETD and what TEI tags are found in which sections.

TEI Tags According to Which Ones are Used in Which ETD Section					
front	div (chapter)	back			
titlePage, docTitle, docAuthor, docDate,	div1, div2, div3,	bibl	appendix	notes "notes" is not TEI	
epigraph, formula, p, author, q, foreign, emph, date, hi	formula, emph, hi, p, author, q, foreign, date	publisher, editor, pubPlace, date, hi, author	argument, p, date, div1, div2, div3, formula, emph, author, q, foreign	formula, emph, p, author, q, foreign, date, publisher, pubPlace, editor, hi	

Table 6: TEI Tags per ETD Section

The "emph" tag is just a TEI-compliant tag for what you might, for instance, wish to emphasize with bold.

While many of the TEI tags are self-evident based on the discussion--for instance of the front matter <u>above</u>, or the sample bibliography entry <u>above</u>--some are especially valuable for organizing your arguments, thesis, and tracking your revisions. **ALWAYS** use this to cite works' titles, NOT the title tag. Since HTML limits "title" to the "head" element, for now, titles other than "docTitle" (the title of your dissertation or thesis), will use worktitle. The "level" attribute is required, signifies whether it is "m" (monographic- a book, monograph, or other publication under a single autonomous title), "s" (a series title), "j" (a journal title), "u" (unpublished title, like a dissertation ;-). Additionally, you can use "class" to specify it is an "abbreviated" version of the title, if it is the "main" title, if it is a "subordinate" title, or, if it is a translation, for instance, it is a "parallel" title. You can choose the format that looks best for you.

The most important TEI tag is the "highlight/"hi" tag (**<hi></hi>**). You will use this throughout your chapters/"div" to mark tehsis themes, revisions, sub-theses, etc. The way you do this is with how well you manage the two attributes of "n" and "id." Think of "id" as a specific designation, and "n" as a category. Thus you can use "hi" as follows to track issues (and it can include almost any matter, including paragraphs and media.

- <hi n="'thesis" id="main"> This sentence addresses the main thesis of the ETD</hi>
- <hi n="thesis" id="sub-argument regarding isomorphing">This would be a sub-thesis regarding isomorphing (I just wanted to use that word).</hi>
- <hi n=''revision'' id=''Prof. Doctor Person demanded this''>Here is the stuff I didn't want to deal with but have to in order to keep one member of my committee happy.</hi>
- <hi n="revision_Prof Person" id="isomorph clarification 2/16/99">Here is another attempt to get my isomorph point past a particularly persnickety committee member.</hi>

Use this also in the front matter with your abstract. **Note**: you can NOT use "hi" in the appendix, you must use "argument" instead.

In addition to the "hi" and "argument" tags, there is the quote/"q" tag. The "q" tag likely would be used, for instance, inside the "blockquote" tag (the blockquote tag is HTML for formatting) in order to give specific information about quotes. It can also be used 'inline" with a single sentence quote too, of course:

<d n="Wittgenstein's magnum opus" id="second edition version"> Here is a
blockquote of text, supposedly from a lofty or knowledgeable tome of wisdom which was
chosen for quoting in the ETD. It is also marked with the "q" tag to show that it is part of the
main argument.

As you can see, the "id" and "n" attributes can be used a variety of ways. Set up your own system for the benefit

Sunday, August 29, 1999

of your study, your committee, and your own reference!

Another important tag is the "formula" tag, for inserting complex materials from mathematical formula's. Similarly, foreign languages are identified with "foreign" tags. The foreign tag accepts an "id" attribute so you can distinguish between several different languages (useful also for non-roman script languages to help browsers read them properly) like the following example:

<foreign id="German">aber nicht hier</foreign>

These tags should be used frequently because they will ensure the future viability, searchability and, ultimately, quotability of your research in the field. In other words, the more you identify of your data, and the more attributes you use systematically, the more accessible and, probably, influential your study will be in your field.

Here is the code for the front file. You may use some sections, or not use others, but this provides a view of what all is usually included. If you are doing a Ph.D., you'll need to add additional committee members to the Certificate of Approval. This file is included on the disk you received from the Graduate College. You are encouraged to use it as a template so as to insure easy and consistent formatting. The many blank lines after the "pre" tags are there to insure formatting.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html SYSTEM "tdm.dtd">
<html>
<head>
<title>
Put your title here.
</title>
<link rel="stylesheet" type="text/css" href="tdm.css"/>
</head>
<body bgcolor="#ffffff">
<front>
<titlePage>
```

<docTitle> <center>

YOUR THESIS TITLE

</center> </docTitle>

<center>

<docAuthor> by

YOU!!! </docAuthor>

</center>

<docDate> A DATE YOU HAVE LONGED FOR </docDate>

Thesis supervisor: <thesisadvisor>Your Mentor</thesisadvisor></center>

</titlePage> <thesiscopyright> <center> Copyright by

YOU **** All Rights Reserved </center>

 <pb/> </thesiscopyright>

<certifapproval> <center>

Graduate College The University of Iowa Iowa City, Iowa

 <center> CERTIFICATE OF APPROVAL Sunday, August 29, 1999

MASTER'S THESIS This is to certify that the Master's thesis of <docAuthor>******** </center> has been approved by the Examining Committee for the thesis requirement for the ******* degree in ****** at the SOME DATE graduation. Thesis Committee: name Thesis supervisor name
 Member

 NAME
 Member
 </center> </certifapproval>

<dedication>

<center>
To very patient people everywhere
</center>

Sunday, August 29, 1999

Page: 28

 </dedication> <epigraph>

</center>

 <pb/> </epigraph> <acknowledgements>

```
<center>
<br /><center>ACKNOWLEDGMENTS
</center>
```

Many people have encouraged and supported me</center>

```
<pb/>>
</acknowledgements>
<toc>
<center>
TABLE OF CONTENTS
</center>
<br />
<a href="back.xml#figures" rel="ref">
LIST OF FIGURES</a>
links to the rest of your ETD
additional links
perhaps a subsection
. .. or two 
and another chapter
another chapter again, and so on
CONCLUSION
<a href="back.xml#notes" rel="ref">
NOTES</a>
<a href="back.xml#bibl" rel="ref">
BIBLIOGRAPHY</a>
<a href="back.xml#append" rel="ref">
APPENDIX</a>
</toc>
<figurelist>
<br />
<br />
<br />
<br />
<br />
<br />
<center><a name="figures" rel="anchor"></a>
<b>LIST OF FIGURES</b></center>
<br />
<01>
"A picture"
```

```
"better than 1000 words""more pictures"</figureList>
```

```
</front>
</body>
</html>
```

Remember that the white space between lines insures proper formatting. It should be preserved unless you are told otherwise.