How to Craft a Basic TEI Header: A Beginner's Guide

In order to conform to TEI standards, a digital text must include a properly formed *header*, which documents the digital text, its source, encoding, and revisions. This guide will provide an introduction to crafting a basic TEI header using [Atom](#) software. For more information, visit the [most recent TEI guidelines](#).

> *Hint:* When you come across an unfamiliar element or tag, use the search bar on the [TEI website](#) to learn more about it and how it can be used in a well-formed encoded document.

Getting Started

The header element opens with a <teiHeader> tag and closes with </teiHeader>. Within the opening and closing tags, headers vary in length and content, depending on how much information an encoder has and/or can document. Even the most basic header, however, must provide a *file description* (<fileDesc>), which contains a *title statement* (<titleStmt>)*, publication statement* (<publicationStmt>)*, and *source description* (<sourceDesc>).*

Refer to the basic TEI encoding template below:

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
   <teiHeader>
```

```
    <fileDesc>
       <titleStmt>
          <title>Title of the digital file</title>
       </titleStmt>
       <publicationStmt>
          <p>Information about publication</p>
       </publicationStmt>
       <sourceDesc>
          <p>Information about the source</p>
       </sourceDesc>
    </fileDesc>
 </teiHeader>
 <text>
    <body>
       <p>Some text here.</p>
    </body>
 </text>
</TEI>
```

The File Description

The file description provides information *about the digital file*. It can contain up to seven different elements, but only three of them are mandatory: a *title statement, publication statement,* and *source description.*

- The first mandatory element nested within the file description is the title statement, contained by <titleStmt> and </titleStmt> tags. The title statement uses additional tags (like <title>, <author>, and <respStmt>) to provide a reader, user, librarian, or archivist with crucial identifying information about the digital file and those responsible for its creation.
- The second mandatory element nested within the file description is the publication statement, contained by <publicationStmt> and </publicationStmt> tags. The publication statement uses additional tags (like <publisher>, <pubPlace>, and <date>) to provide a reader, user,

librarian, or archivist with information about the digital file's publication and distribution.

- The third and final mandatory element nested within the file description is the source description, contained by <sourceDesc> and </sourceDesc> tags. The source description uses additional tags (like <bibl>, <biblFull>, <biblStruct>, or <msDesc>) to provide a reader, user, librarian, or archivist with more information about the *source from which a digital file is derived or created.*

Crafting a Basic Header

To craft a basic header, simply populate the mandatory tags with the appropriate information for your digital text. As a model, we'll craft a basic header for an article entitled, "Nation moved by magnificent march of protest: The march," which appeared in a 1963 issue of the *Baltimore Afro-American* newspaper.

First, **open a new file in Atom** and begin with a standard TEI template:

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
     <fileDesc>
        <titleStmt>
           <title>Title of the digital file</title>
        </titleStmt>
        <publicationStmt>
```

```
            <p>Information about publication</p>
         </publicationStmt>
         <sourceDesc>
             <p>Information about the source</p>
         </sourceDesc>
      </fileDesc>
   </teiHeader>
   <text>
      <body>
         <p>Some text here.</p>
      </body>
   </text>
</TEI>
```

Then, **save the file with an .xml extension** to activate your TEI validation plug-in. As you continue to work, try to save often.

Begin with the first mandatory element, the title statement. If you visit the <titleStmt> element in the TEI P5 guidelines, you'll see that the <titleStmt> tag can contain <title>, <author>, and <respStmt> tags to thoroughly identify the digital file we are about to create.

```
<titleStmt>
   <title>Title of the digital file</title>
   <author>Author of the digital file</author>
   <respStmt>
      <resp>A phrase describing the nature of a person's
      role or responsibility to the digital file</resp>
      <persName>Name of the contributor</persName>
   </respStmt>
</titleStmt>
```

Once you have identified the tags contained by the <titleStmt> element, you can populate them with the appropriate information for the selected article.

```
<titleStmt>
   <title>Nation moved by a magnificent march
   of protest</title>
   <author>
     <persName>Jessica H. Lu</persName>
   </author>
   <respStmt>
      <resp>Encoded by</resp>
      <name>Jessica H. Lu</name>
   </respStmt>
</titleStmt>
```

The above encoding represents a basic effort to provide title information.
However, we can be even more thorough by utilizing attributes, identifiers, and
additional tags. For example:

- The original headline, "Nation moved by magnificent march of protest,"
  from the *Baltimore Afro-American* housed *two* articles: "The march" and
  "The color." To clarify which one is encoded in this digital file, we can use
  the *type* attribute ("type=") to specify the *main* title from the *sub-title.*

- The *title* should be the title *of the digital file,* rather than the newspaper
  article from which the digital file is derived. So, we can add an additional
  sub-title to identify the file as the *digital edition*, distinct from the original
  newspaper copy*.*

- When the name appears within the <author> tag, we can assign it a unique
  identifier (an "xml:id"). Then, later in the digital file, we can *reference* that
  unique identifier when that person's name appears again.

- Since the author of this digital file is associated with an organization, who
  may share part ownership of the digital file, we can add the name of that
  organization using an <orgName> tag. Additionally, we can assign the
  organization a unique identifier, just as we did for the author.

```
<titleStmt>
   <title type="main">Nation moved by a magnificent march
   of protest</title>
   <title type="sub">The march</title>
   <title type="sub">Digital Edition</title>
   <author>
     <persName xml:id="JLU">Jessica H. Lu</persName>
     <orgName xml:id="AADHUM">African American History,
     Culture, and Digital Humanities (AADHum)</orgName>
   </author>
   <respStmt>
     <resp>Encoded by</resp>
     <persName ref="#JLU">Jessica H. Lu</persName>
   </respStmt>
</titleStmt>
```

Once you've populated the <titleStmt> element, **verify that your TEI is well-formed.** Take a look at your Atom screen. If a red dot appears along the left-hand sidebar, you may have made an error. Use your mouse to hover over the erroneous line of markup; a pop-up will appear to help you identify the problem.

Now, **you can move on to the next mandatory element: the publication statement.** If you visit the <publicationStmt> element in the TEI P5 guidelines, you'll see that the <publicationStmt> tag can contain <publisher>, <pubPlace>, and <date> tags to provide publication information about the digital file we are about to create.

```
<publicationStmt>
   <publisher>Name of the publisher</publisher>
   <pubPlace>Location of the publisher</pubPlace>
   <date>Publication date</date>
</publicationStmt>
```

Once you have identified the tags contained by the <publicationStmt> element, you can populate them with the appropriate information for the selected article.

```
<publicationStmt>
   <publisher>
      African American History, Culture, and Digital
      Humanities (AADHum)
   </publisher>
   <pubPlace>College Park, MD</pubPlace>
   <date>October 16, 2017</date>
</publicationStmt>
```

The above encoding represents a basic effort to provide publication information. However, we can once again be even more thorough. For example:

- We can use the <orgName> tag to specify that the publisher is an organization.
- Since we assigned this organization a unique identifier earlier in the header, we can *reference* that identifier here.
- We can use additional tags to identify the two parts of the publication location: the city and the state. In TEI, the city is identified by a <settlement> tag, whereas the state is identified by a <region> tag.
- We can help a computer read the date by using a *when* attribute ("when=") to format it in a machine-readable (e.g. *numeric*) way: by year, month, and day.

```
<publicationStmt>
   <publisher>
      <orgName ref="#AADHUM">African American History,
         Culture, and Digital Humanities (AADHum)</orgName>
   </publisher>
   <pubPlace>
      <settlement>College Park</settlement>
      <region>Maryland</region>
   </pubPlace>
```

```
   <date when="2017-10-16">October 16, 2017</date>
</publicationStmt>
```

We're two-thirds of the way there to a complete, basic header! Make sure you've saved your document and have verified that your markup has been well-formed before moving forward.

The final mandatory element of the file description is the source description. If you visit the <sourceDesc> element in the TEI P5 guidelines, you'll see that the <sourceDesc> tag can contain several different tags, like <bibl>, <biblFull>, <biblStruct>, and <msDesc>. Each of these tags contains different type of bibliographic information, so you can choose the one that is most appropriate. For this basic header, we will use the simplest one: the <bibl> tag. <Bibl> allows you to provide loosely-organized bibliographic information, the subcomponents of which may or may not be explicitly tagged, *about the source from which the digital file is derived or created* (in this case, this particular issue of the *Baltimore Afro-American*).

```
<sourceDesc>
   <bibl>
   </bibl>
</sourceDesc>
```

Since the <bibl> tag is so flexible, you can populate it with the information you know you have. Refer back to the masthead of the *Baltimore Afro-American* to help you identify, for instance, the <title>, <publisher>, <pubPlace>, and <date>.

```
<sourceDesc>
   <bibl>
      <title>Baltimore Afro-American</title>
      <publisher>Afro-American Co.</publisher>
```

```
        <pubPlace>Baltimore, MD</pubPlace>
        <date>August 31, 1963</date>
    </bibl>
</sourceDesc>
```

As before, when we further developed our *publication statement*, we can provide additional information to help the computer (and other users) read this bibliographic information.

```
<sourceDesc>
    <bibl>
        <title>Baltimore Afro-American</title>
        <publisher><orgName>Afro-American
        Co.</orgName></publisher>
        <pubPlace>
            <settlement>Baltimore</settlement>
            <region>MD</region>
        </pubPlace>
        <date when="1963-08-31">August 31, 1963</date>
    </bibl>
</sourceDesc>
```

(Have you saved your xml file recently?)

Congratulations! Having completed your source description, you are now finished with a basic header! Your markup should look something like this:

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://relaxng.org/ns/structure/1.0"?>
<?xml-model
href="http://www.tei-c.org/Vault/P5/current/xml/tei/custom/
schema/relaxng/tei_all.rng" type="application/xml"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
```

```
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <teiHeader>
      <fileDesc>
          <titleStmt>
              <title type="main">Nation moved by magnificent
              march of protest</title>
              <title type="sub">The march</title>
              <title type="sub">Digital Edition</title>
              <author>
                  <persName xml:id="JLU">Jessica H.
                  Lu</persName>
                  <orgName xml:id="AADHUM">African American
                  History, Culture, and Digital Humanities
                  (AADHum)</orgName>
              </author>
              <respStmt>
                  <resp>Encoded by</resp>
                  <persName ref="#JLU">Jessica H.
                  Lu</persName>
              </respStmt>
          </titleStmt>
          <publicationStmt>
              <publisher>
                  <orgName ref="#AADHUM">African American
                  History, Culture, and Digital Humanities
                  (AADHum)</orgName>
              </publisher>
              <pubPlace>
                  <settlement>College Park</settlement>
                  <region>Maryland</region>
              </pubPlace>
              <date when="2017-10-16">October 16,
              2017</date>
          </publicationStmt>
          <sourceDesc>
              <bibl>
                  <title>Baltimore Afro-American</title>
                  <publisher><orgName>Afro-American
                  Co.</orgName></publisher>
                  <pubPlace>
```

```
                         <settlement>Baltimore</settlement>
                         <region>MD</region>
                   </pubPlace>
                   <date when="1963-08-31">August 31,
                   1963</date>
             </bibl>
       </sourceDesc>
    </fileDesc>
  </teiHeader>
  <text>
     <body>
        <p>Some text here.</p>
     </body>
  </text>
</TEI>
```

Let's practice!

Choose one of the articles from the [1967 *The Diamondback* series on African American life at the University of Maryland](#). Following the same model detailed above, craft a basic header for your selected article!

1. January 5: "The negro on campus"
2. January 6: "Negroes on campus: Integration at College Park"
3. January 9: "Negroes on white campus: Apprehension limits racial interaction"
4. January 9: "'Average' freshman: 'Class of '70' described"
5. January 11: "The negro on campus: Racial issues down since 1954"
6. January 12: "A lonely walk?: Negroes view campus with mixed emotions"
7. January 13: "Interdating: The negro point of view"

Ready for more?

Once you feel comfortable with this basic header structure, you can begin moving toward more advanced headers. In addition to the *file description*'s

(<fileDesc>) three mandatory elements covered here (<titleStmt>, <publicationStmt>, and <sourceDesc>), there are four others you can consider incorporating:

- *Edition statement* (<editionStmt>): provides information about the edition of the digital text.
- *Extent* (<extent>): provides information about size (as stored on some carrier medium, whether digital or non-digital) and scope of the digital text. This is particularly useful when the digital text marks up only a portion of the original source from which the digital text is derived.
- *Notes statement* (<notesStmt>): for any notes providing additional information about a text otherwise unaddressed by the other file description elements.
- *Series statement* (<seriesStmt>): groups information about the series, if any, to which a text belongs.

In addition, a header may also include other elements *in addition to, but not contained by,* the file description. For example:

- *Encoding description* (<encodingDesc>): describes the relationship between the digital file and the original source from which it is derived.
- *Profile description* (<profileDesc>): describes non-bibliographic aspects of a text, especially the languages and sublanguages used, as well as the situation in which it was produced, the participants, and their setting.
- *Revision description* (<revisionDesc>): records information about the revisions to the digital file.

Finally, keeping in mind that you'll be exploring the possibilities of using TEI to conduct a digital analysis of *ideas,* consider ways to set up that analysis within the header of the digital text. ***Can you search the guidelines for helpful elements?***