

Getting acquainted with oXygen

Lou Burnard Consulting

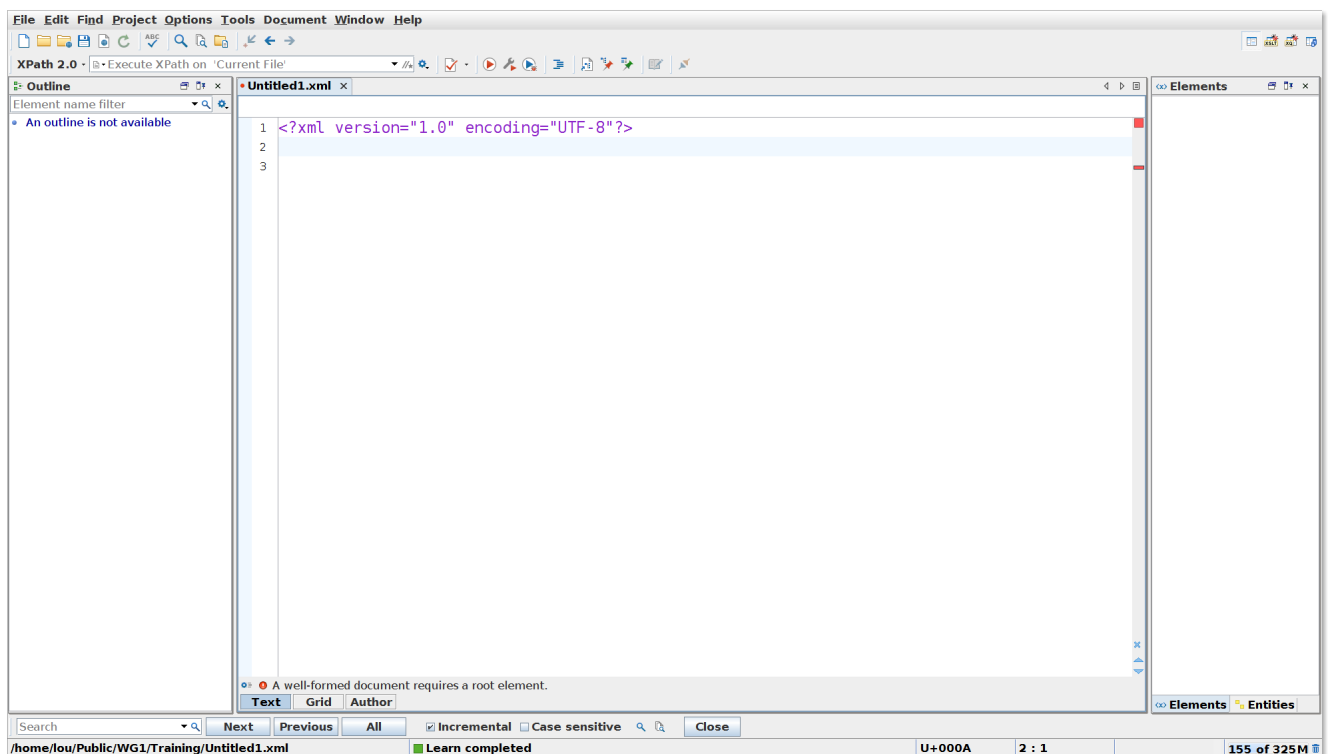
2018-04-14

This first tutorial explores the basic features provided by oXygen for the creation of a well formed and valid XML documents. It has three parts:

1. introducing tags into a well formed document
2. using a simple TEI schema
3. displaying tagged text

1 My first XML document

- Start up oXygen.
- Click on the New icon (first left on the toolbar); or select File->New from the menu; or type CTRL-N
- In the New dialogue, select New Document - Document XML and click the Create button.
- You are now in the typical oXygen workspace. There is a large white window in the middle where you are going to work on your document, with some smaller satellite windows we will look at later.



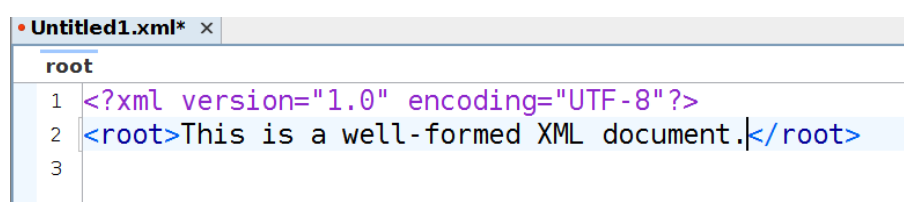
- Your workspace currently contains just one line: an XML Declaration like this:

```
<?xml version="1.0" encoding="UTF-8"?>
```

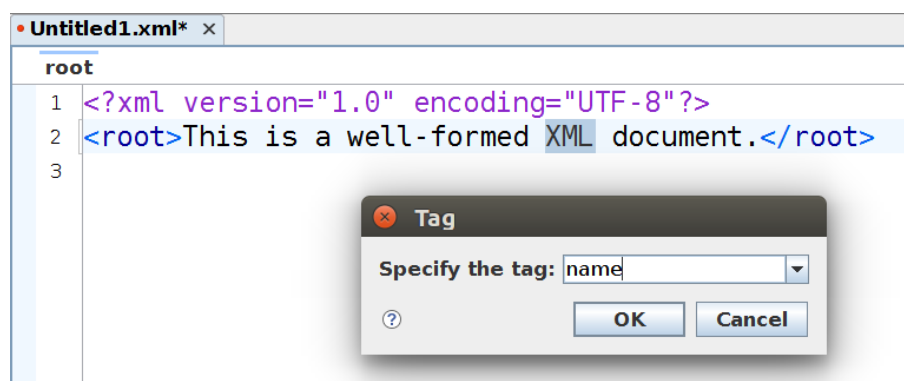
1 MY FIRST XML DOCUMENT

Every XML document has to begin with a line like this. It shows a processor that this is an XML file, which version of XML this file is using, and what character encoding has been used for its textual content. XML version 1.0 is a W3C recommendation from 2008. UTF-8 (Universal Character Set Transformation Format - 8 bit) is the usual way of representing the Unicode character set, which supports most characters from nearly all human writing systems.

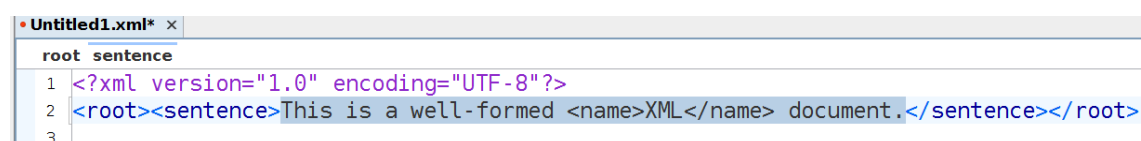
- The little red square at top right shows we have a problem! Look at the error message at the bottom of the screen: ‘A well-formed document requires a root element’. Ah yes, an XML document must have at least one tag in it. Type the start-tag `<root>` oXygen obligingly completes this with a closing tag `</root>` as soon as you hit the `>` character. You now have a well-formed XML document and the little red square becomes green. Well done! Let’s make the document a bit more interesting by adding some content to the element, for example, type ‘This is a well formed XML document.’ between the start and end tags.



- Oxygen provides several ways of tagging documents. Let’s try a second one. Select the word XML with the mouse and then type CTRL-E. Now you can enter a name for the tag you’re adding: we suggest name Click OK to continue

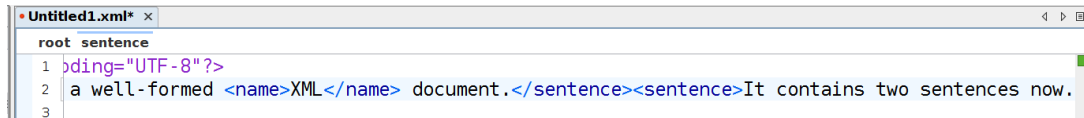


- XML elements can nest within each other: we have a `<root>` element which contains a bit of text, followed by a `<name>` element, followed by another bit of text. Select all three of these (i.e. everything inside the `<root>`) and tag it `<sentence>` again using CTRL-E.



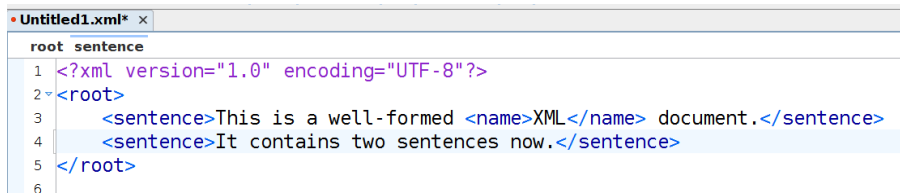
- The third way of tagging a document is to divide up its elements automatically. Put the cursor at the end of the `<sentence>` element you just created (between `document.` and `</sentence>`). Type ALT-SHIFT-d to close this `<sentence>` and start a new

one. You can now add a second sentence to your document. The screen will scroll to the right as you type:



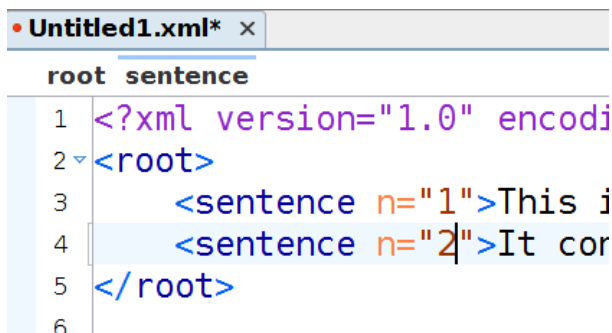
```
Untitled1.xml* x
root sentence
1 <?xml version="UTF-8"?>
2 a well-formed <name>XML</name> document.</sentence><sentence>It contains two sentences now.
3
```

- To make the document look a bit nicer, type CTRL-SHIFT-p or click the Format and Indent button.



```
Untitled1.xml* x
root sentence
1 <?xml version="1.0" encoding="UTF-8"?>
2 <root>
3   <sentence>This is a well-formed <name>XML</name> document.</sentence>
4   <sentence>It contains two sentences now.</sentence>
5 </root>
6
```

- oXygen also helps you when it comes to adding attributes. Put the cursor *inside* the first `<sentence>` tag, just in front of the `>` and type a space, followed by `n="` . Oxygen adds a `"` and the cursor stays there, waiting for your input. Type the figure 1.
- Number the second sentence "2" in the same way.



```
Untitled1.xml* x
root sentence
1 <?xml version="1.0" encoding="UTF-8"?>
2 <root>
3   <sentence n="1">This is a well-formed <name>XML</name> document.</sentence>
4   <sentence n="2">It contains two sentences now.</sentence>
5 </root>
6
```

- Oxygen does its best to keep your document well-formed. Type a start-tag for a new element somewhere (`<wibble>` for example) and oXygen will immediately add a corresponding end-tag (`</wibble>`). If you change the name inside a start-tag, the end-tag will change to match it. However, if you change the name inside an end-tag, oXygen signals an error. Type CTRL-z to undo your last change to the document at any point.

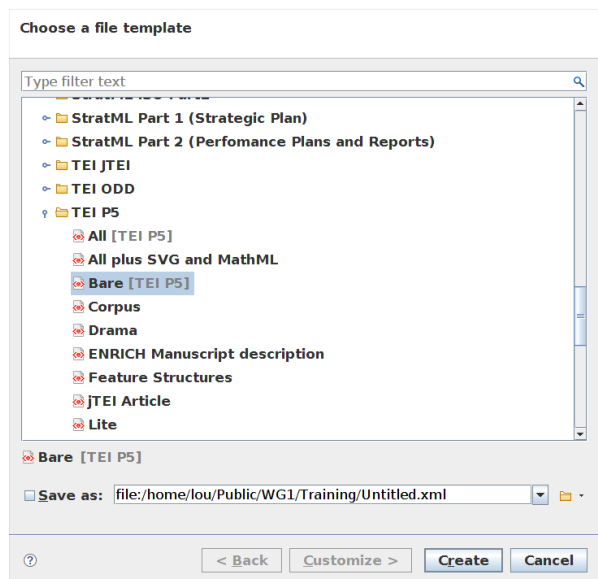
What have we learned? There are three ways of adding tags to a document :

- if you type a start-tag in the right place, oXygen will add the corresponding end-tag, and leave the cursor in the right place for you to add some content
- if you highlight a bit of text, you can type CTRL-E to enclose it with start and end tags
- if you put the cursor anywhere inside an element and type ALT-SHIFT-d, oXygen will divide the element in two at that point
- and fourthly, you add attributes and their values inside a start-tag

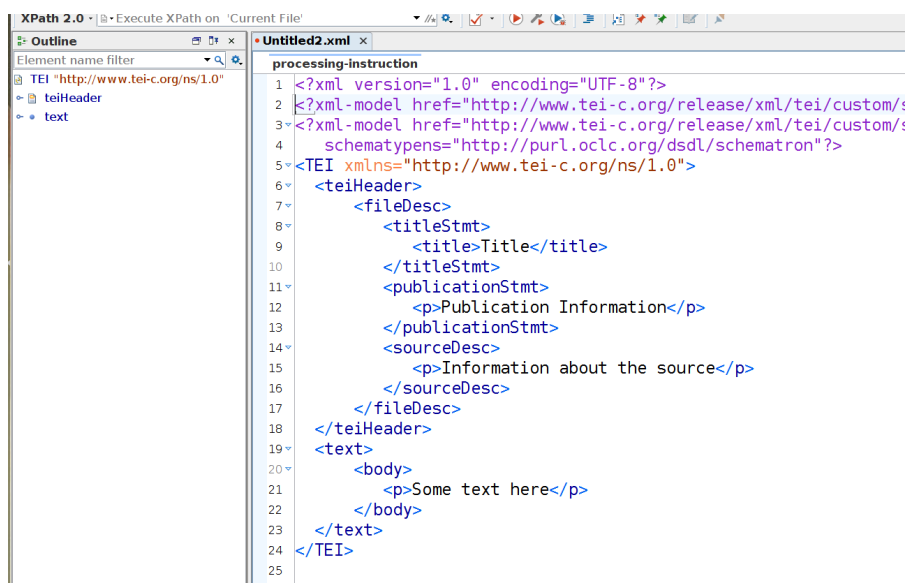
2 Using a schema

So far, oXygen didn't do anything to control which tags you used in your document. It just ensured they were syntactically correct and that the document was well formed. In this exercise you will use a predefined XML vocabulary to control which tags and attributes are available in your document.

- Open oXygen and click on the New icon (first left on the toolbar); or select File->New from the menu; or type CTRL-N.
- The New File dialogue appears. Scroll down to see some of the options under New Document: oXygen can work with all kinds of files, not only XML. Further down, there is an option to choose from a long menu of pre-defined XML vocabularies. oXygen comes ready to work with many of these, including several varieties of TEI. As yet there is no ELTeC schema here ; instead we will use a minimal TEI schema called TEI Bare.
- In the New dialog, scroll down to the section Framework templates, within that select TEI P5, and within that select Bare.



When you press Create this time, oXygen displays the basic structure of a TEI document, ready for you to complete. In the side window labelled Outline (select Window->Show View-Summary if this is not visible) you can see that a document TEI has two major parts: a `<teiHeader>` element containing metadata, which is followed by a `<text>` element, containing the text itself.



2.1 Add the metadata

Every TEI document requires at least three pieces of metadata: a title, information about its publication, and information about its source. For this exercise, we suggest:

- Type a title for the document you're creating between the `<title>` and `</title>` tags: 'My second XML document' for example.
- Between the `<p>` and `</p>` tags following the `<publicationStmt>` tag type in a phrase such as 'Unpublished training exercise'.
- Between the `<p>` and `</p>` tags following the `<sourceDesc>` tag type in a phrase such as 'No pre-existing source'.

Because you are using a schema, oXygen will only allow you to add elements defined by that schema.

- Put the cursor after the `</title>` end-tag and before the `</titleSmt>` end-tag.
- Type `<` and wait for a moment.
- The green square becomes red, because your document now has a `<` which is not followed by the name of an element.
- oXygen suggests a list of elements that you could legally insert at this point in the document: for example, you could add an `<author>` element, or a second `<title>`, amongst other things.
- hit the ENTER key to accept the first suggestion. oXygen inserts start- and end-tags for an `<author>` element.
- The cursor stays inside the start-tag for the element just in case you wanted to add some attributes for the author. Type a space to see what attributes are available. You don't need any of these, so type a backspace and move the cursor to the right so that you can type in your name as content of the `<element>`, since you are the author of this document.

2.2 Adding text to the document

- Select everything inside the `<body>` element, i.e. the sequence `<p>Some text here</p>`, including the tags.
- Type **CTRL-e** (or select Document->Markup-> Surround with tags from the menu).
- As before, oXygen shows a list of the tags which are legal at this point; in this case, you should choose `<div>`, and then click on OK.
- This `<div>` element will contain the first chapter of your novel. Put the cursor inside the `<div>` start-tag and type a space to see what attributes are available. This time, select the `@n` attribute which you can use this to number the chapter. Hit RETURN to insert it into your document, and give it a value **1**.
- Repeat this manoeuvre to add a `@type` attribute with the value **chapter**. If you'd prefer to work in some language other than English, you could also add an `@xml:lang` attribute to indicate the language of your chapter. (Languages are indicated using ISO 639-2 language codes e.g. **fr** for French, **sb** for Serbian etc.)
- The first chapter should have a title such as 'Chapter One' or 'In the beginning', or whatever you fancy. Add this title as the content of a `<head>` element inserted at the beginning of your `<div>`.

Your document should now look something like this :

```
<text>
  <body>
    <div type="chapter" n="1">
      <head>Chapter One</head>
      <p>Some text here</p>
    </div>
  </body>
</text>
```

Over to you to complete this chapter, preferably with some text a little more interesting than 'some text here' ! Describe your first reactions to using oXygen, or your humble origins and early career as a researcher; share your ideas on the current political situation; write a romantic novel; whatever you like...

Of course, you'll need more than one paragraph to do your ideas justice. Do you remember from the first exercise how to split this paragraph and begin a new one? Assuming you now have several hundred words in your first paragraph and you need to split it up:

- Put the cursor at the point in your text where each new paragraph should begin.
- Type **ALT-SHIFT-d** (or Document -> Markup -> Split Element).

If you want to start a new chapter, what will you do?

3 Displaying tagged text

Is the amount of tagging in your document beginning to worry you a little?

You can tidy up the way the tagged text is displayed, as you know (click on the Indent button, select Document-> Source->Format and Indent, or **CTRL-SHIFT-P**). But maybe you'd rather see fewer tags?

- At the bottom of the screen there are three buttons marked Text, Grid, and Author. Select the last one. The way your document is displayed changes and a new TEI P5 menu

appears above the toolbar. Choose **Tags display mode** from this menu and experiment with some of the possibilities it offers as regards display of the tags in your document.

- In any case, you can see that your document is now formatted : headings are displayed in a different size, line breaks are treated as white space, paragraphs are tidily aligned and so on.
- Note that you are still editing the document, even though it looks different. If you make a change in Author mode (delete or add a word, for example) and then switch back to Text mode, the change will appear there as well.