



The Open University

---

openlearn

---

## Hints and tips: Using OpenLearn XML

1. Overview
2. Your starting point
3. Generic editorial set up and brief
4. Metadata
5. Basic XML tag structure



## Overview

---

What follows below is a detailed description of how our XML schema has been used to generate learning content for the OpenLearn initiative. Following the hints below should help you rework our content successfully and share it with the OpenLearn community.

## Your starting point

---

The file you start from already works. We know it does because we are delivering it through to the web. If you do nothing to the XML tags at all, but change content within the tags, you can be fairly sure your unit will validate too and will successfully publish to the Labspace. Don't be discouraged from being more ambitious, but don't be surprised if we can't display your work if it doesn't fit with the underlying XML structures of our learning materials.

All initial materials processed for delivery through the OpenLearn Learningspace and Labspace are tagged within a generic production schema. To rework these units and resubmit them for delivery through to the OpenLearn Labspace, you will need to work in the XML file itself and validate it before upload.

Reworked versions of materials will be shown alongside other versions of the unit, so users can move through the content as learners and can see the differences made to the version they are working on (such as additional case studies, examples).

If you are planning to retain the materials embedded within the unit (such as images, video, pdfs etc) as they are, then you do not need to change the tags which refer to those. If you plan to introduce other elements, these need to be referred to according to the conventions of the rest of the document.



## Generic editorial set up and brief

---

The original XML files have come from materials originally held in a variety of formats and have all been 'scrubbed' and tagged up according to a consistent production schema (and rewritten to a certain extent to make sense on the web). To edit XML files we suggest you use one of the many XML editing tools freely available. We have used amongst other things Word 2003 Professional Edition, XMLSpy Home Edition, and other tagging tools in which it's simple to add and edit tags and attributes on tags. Whichever you choose, the tool should also quickly and easily validate your files, and give helpful information about the location and nature of any problems that prevent validation.

### Set up

All XML documents must validate against the OU Generic Production schema. This is supplied to you when you download a unit. In order to validate XML files they must be stored in the same folder on your PC as the schema. We recommend you follow the approach set out below (if you are a PC user):

To create a new folder

1. Open Windows Explorer.
2. Click the drive or folder in which you want to create a new folder. This should be **C**
3. On the **File** menu, point to **New**, and then click **Folder**.
4. Type a name for the new folder (this should be OCI), and then press ENTER.

### Associating the schema

On opening an XML file first associate the schema. This is done in different ways with different tools. In XMLSpy, for example, go to DTD/Schema → Assign schema. You'll get a dialogue box asking you if you really want to proceed: you do, click ok. You'll then get another dialogue box in which you can browse to pick up the schema. This will be in the same folder. Find it and click ok.

Before doing anything else make sure that the in the <item> tag on the second line of the file the value on the second attribute reads only "OUGenericFull.xsd". Remove any extraneous filepath information (like 'C:\OCI\') if necessary.



### Validating the XML file

The first thing to do once the file is open and the schema associated is to ensure that the XML file validates against the schema: in XMLSpy for example you would go to XML → validate (or click on the icon with document and green tick, or press F8). In the small window at the bottom of the screen you should get a green tick icon and a message saying 'File C:\oci\filename.xml is valid.'

If the file is invalid, most XML editors will identify the point where the code is invalid, and tell you what's wrong: e.g., missing end tags; wrong attributes etc.

### The 'integrity model'

Units originally produced for initial delivery through OpenLearn have been worked on under what we've been calling 'integrity model'. This entails that:

*all the material in the Unit is recognisably very similar to the original material and as complete as possible with the ability to study it in exactly the same order as the original. In practice, this means slicing existing text into smaller portions, each of which being delivered as a text page or pdf, with activities translated into an interactive form, and other resources such as audio clips added to it rather than being delivered separately.*

Andy Lane, OCI project director

### Sessions, Sections, SubSections, SubSubsections

Each OCI unit should consist of one set of <Session> tags encompassing the main body of the learning material. The session tag takes a numbered id attribute which, as there will always be one session per unit, will be id="SES001": i.e. <Session id="SES001">.

Within a Session <Section>, <SubSection> and <SubSubSection> tags used to \*embody\* the structural elements of the unit. It's important to note that these tags must be properly nested, which is to say that a SubSubSection must be within a pair of SubSection tags, which must be in Section, which must in turn be in Session.

Section and SubSection tags are used to generate screen breaks. The way we have decided units will be rendered into HTML means that all content must be contained within one 'level' of tags, either Section or, more likely, SubSection. Should any material be placed between the start of a Section and start of a SubSection it will be lost in the transition from XML to HTML.

Section and SubSection tags also take hierarchical, sequential attributes: i.e., the first Section tag in a Session is <Section id="SEC001\_001">.



SubSubSection tags follow the same pattern, but they don't create any screen breaks, so are used for breaking material up within screens. A 'collapsed' structure diagram of these elements is below:

```
<Session id="SES001">
  <Section id="SEC001_001">
    <SubSection id="SEC001_001_001">
    </SubSection>
    <SubSection id="SEC001_001_002">
      <SubSubSection id="SEC001_001_002_001">
      </SubSubSection>
    </SubSection>
    <SubSection id="SEC001_001_003">
    </SubSection>
  </Section>
  <Section id="SEC001_002">
    <SubSection id="SEC001_002_001">
    </SubSection>
    <SubSection id="SEC001_002_002">
    </SubSection>
  </Section>
  <Section id="SEC001_003">
  </Section>
</Session>
```

Each pair of Session, Section and SubSection tags requires a pair of Title tags immediately after their respective start tags while SubSubSection requires Heading tags.

For a guide to which tags can legitimately sit within Section, SubSection, and SubSubSection, see the section on XML tag structure below.



## Front matter: introduction and learning outcomes

Each unit requires an introduction of up to 200 words which will be pulled through onto the course contents page where users will be browsing or searching for courses. It should provide a brief guide to the unit, and should complement the unit's learning outcomes (see below). This introduction will be part of the 'front matter' of the unit, rather than in the main body of the text, and may sit alongside and even be cannibalised from an introduction in the main body.

Each unit requires learning outcomes: usually a bulleted list of points. The list will on average contain three or four items but may contain many more, or as few as one. For our own materials, front matter has been structured as below:

```
<FrontMatter>
  <CourseCode>e.g. T173</CourseCode>
  <CourseTitle>full course title </CourseTitle>
  <ItemID> </ItemID>
  <ItemTitle>originating book and block title</ItemTitle>
  <Imprint>
    <Standard>
      <GeneralInfo>
      </GeneralInfo>
      <Address>
        <AddressLine></AddressLine>
      </Address>
      <FirstPublished>Year</FirstPublished>
      <Copyright>&#xA9;</Copyright>
      <Rights>
        <Paragraph> </Paragraph>
      </Rights>
      <Edited> </Edited>
      <Typeset> </Typeset>
      <Printed> </Printed>
      <ISBN></ISBN>
      <Edition>Second edition</Edition>
    </Standard>
  </Imprint>
  <Introduction>
    <Title></Title>
    <Paragraph></Paragraph>
  </Introduction>
  <LearningOutcomes>
    <Paragraph></Paragraph>
    <LearningOutcome></LearningOutcome>
    <LearningOutcome></LearningOutcome>
  </LearningOutcomes>
</FrontMatter>
```



## Activities and SAQs

Please ensure that all activities and self-assessment questions (SAQs) continue to be appropriate for an **independent** online user. If there are any references to, e.g., tutors, group discussions, or other types of what's called 'supported open learning' (SOL) please either try to rework the activity to remove this aspect, or accept that independent learners will be unable to undertake some of what you suggest.

The opening Activity /SAQ tag has a numeric id attribute, i.e. `<Activity id="ACT001">`. These attributes should be sequential throughout the unit. They need not, however, start from number one.

If the activities headings are numbered, please ensure that these are sequential throughout the unit, and ensure that *they do start at number one*, i.e.:

```
<Activity id="ACT001">  
  <Heading>Activity 1: A question for you to think about</Heading>
```

Please note the styling of the activity heading, and apply this to SAQs and exercises.



Note that Activity tags may contain either Discussion or Answer tags, *but not both*. If you create an activity with answers and discussion, try to find some way of incorporating both within either Answer or Discussion. Activities are by far the most common form of interactive question, but structurally activities and SAQs are virtually identical, though Discussion is disallowed in SAQ. A typical activity is structured as follows:

```
<Activity id="ACT001">
  [<Heading>A heading is not necessary</Heading>]
  <Question>
    [<SubHeading>Not necessary</SubHeading>]
    <Paragraph></Paragraph>
  </Question>
  <Discussion>
    [<SubHeading>Not necessary</SubHeading>]
    <Paragraph></Paragraph>
  </Discussion>
</Activity>
```

### References to other courses/units, and other course material

Please avoid references to other courses (whether from the UK Open University or other providers), or other units within the course from which the unit is drawn. Again this is because materials should be able to stand alone as far as possible.

You may wish to 'hang' materials as pdfs alongside the main text. Have a look for other areas of the unit where this has been done as there is a set of tags which allow you to do this (similar to figure and media content tags). In attaching such material it is your responsibility to ensure you have permission to do so (if it does not belong to you).

We are currently unable to accommodate additional file types included alongside your XML document **apart from**

- Video files (in mpeg4 or .flv format)
- Audio files (in mp3 format)
- Still images (in .jpg format)
- Animations (flash)

To meet the needs of students with disabilities, transcripts for audio and video files and descriptions of images and animations should be supplied with the source materials.





### Other illustrative and textual elements

Other xml elements such as figures, tables, quotes, extracts, boxes and case studies all take sequential attributes on their initial tags in the same manner as activities. Again these sequential attributes need not start at one, but where the numbers are displayed in headings or captions they must start from one at the beginning of the unit.

Examples of these elements are given below (see *Examples of tags in use*)

Styling of table headings is as follows:

```
<TableHead>Table 1: Table showing distribution of data</TableHead>
```

Styling of figure captions is as follows:

```
<Caption><b>Figure 1: Image of interesting item</b></Caption>
```

In many cases the placing of these elements has been determined by traditional page layout concerns. For our purposes, though, elements should be placed as close to their introduction in the text as possible (bearing in mind that in general these sorts of elements are only allowed in Section, SubSection and SubSubSection) without breaking the flow of the text.

### Footnotes and sidenotes

Some of these remain based on the origins of the materials, but footnotes and sidenotes will not work online. Where they do not add anything to the body text they can safely be deleted: otherwise, please find a way of incorporating them into the text.

### Links to external website

Please check and verify links to external websites, and add a note to the text stating when the link was checked. Links to external websites should open a new window in the users web browser. In order that this happens, please check external links for a target="\_blank" attribute. For example:

This information is adapted from Disability Rights Commission, definition of disability at <http://www.drc-gb.org/ knowyourrights/definition.asp>, accessed 24 May 2006.



## Metadata

---

Alongside the main XML file, there is another XML file you will need to complete which captures metadata describing the unit. You will see from the file when you open it that it has already had some fields completed. Some of these can be overwritten, some must be left unchanged, most of it is optional:

Tags	Change	Comment
<identifier>eg A102_1</identifier>	No	Do not change this: it is the code for the original unit
<title>This unit deals with ...</title>	Yes	This may be changed: it is the displayed title of the unit
<description>	No	This field is overwritten when you upload your unit
<classification keywords>	No	Do not change this
<version>	No	This field is overwritten when you upload your unit
<level>	Yes	This may be changed: it is the academic level of the material (but unless you have simplified or made material much more complex it will usually stay the same)
<time>	Yes	This is the time needed to study the unit



## Links to relevant websites

You can also use the metadata to add links within the related resources block. We have included links to the courses the materials come from. You can add to this if you follow this approach:

**In all cases the `http://` must be included.**

```
<relations>
  <relation>
    <urltitle>An Open University course on this topic</urltitle>
    <urladdress>http://www3.open.ac.uk/courses/bin/p12.dll?C01E123</urladdress>
    <urlgroup>OU Links</urlgroup>
  </relation>
  <relation>
    <urltitle>Studying this topic at the Open University</urltitle>
    <urladdress>http://www3.open.ac.uk/courses/classifications/education_and_teacher_training.shtml</urladdress>
    <urlgroup>OU Links</urlgroup>
  </relation>
</relations>
```

## Uploading your file

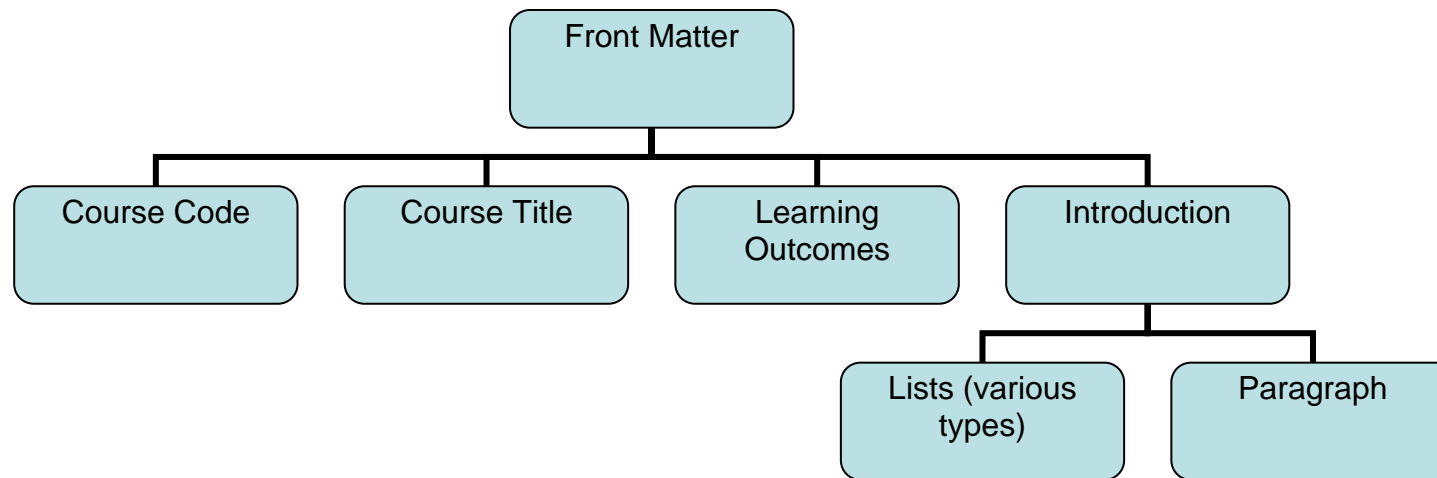
Follow the instructions in the upload area of the Labspace website to send your version of the unit to the site. Provided it passes through the technical checks we apply, it should appear to the OpenLearn community very soon after.

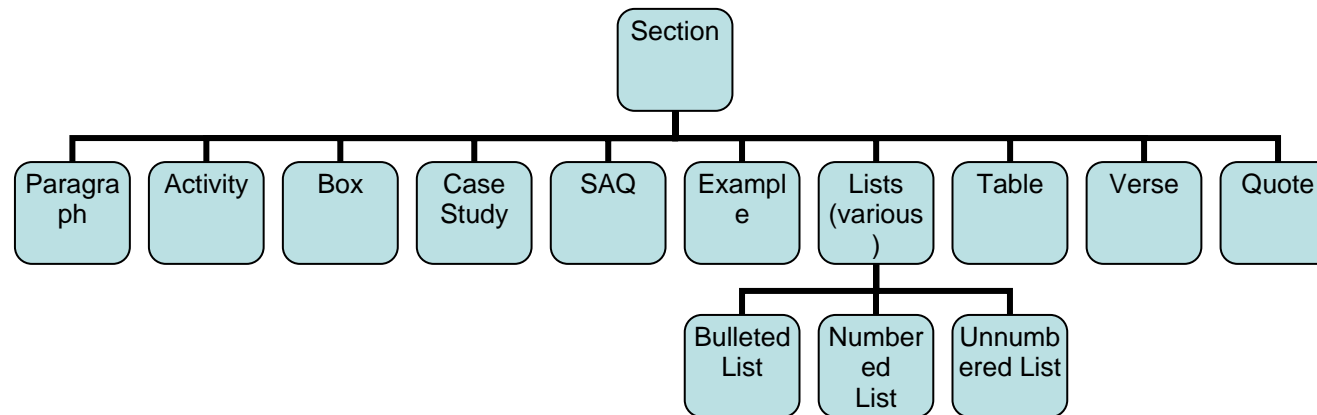


## Basic XML tag structure in OU generic schema

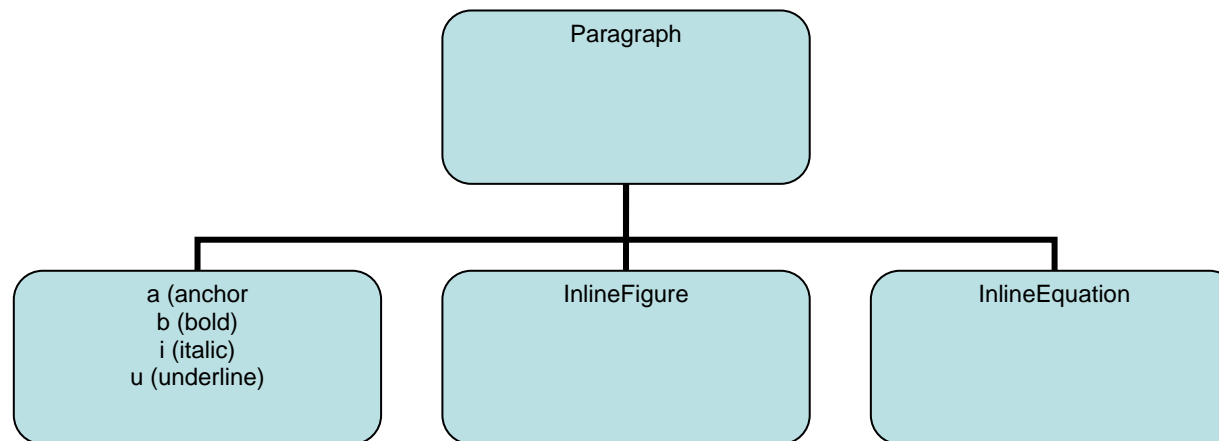
---

This is a brief guide to the tags you can expect to find in one of our xml files and some of the tags you can expect to find within other tags. This list is not exhaustive but covers most of the tags you can expect to see regularly. As far as possible I've tried to list them in the order in which you might expect to see them. A note on the Session tag: Each OCI unit should contain only one set of Session tags which should encompass everything between FrontMatter and BackMatter. The Session tag should be followed immediately by Title and thereafter by Section. Although it is valid XML to include other other elements in Session it is OCI house style include **only Section** and **Title**

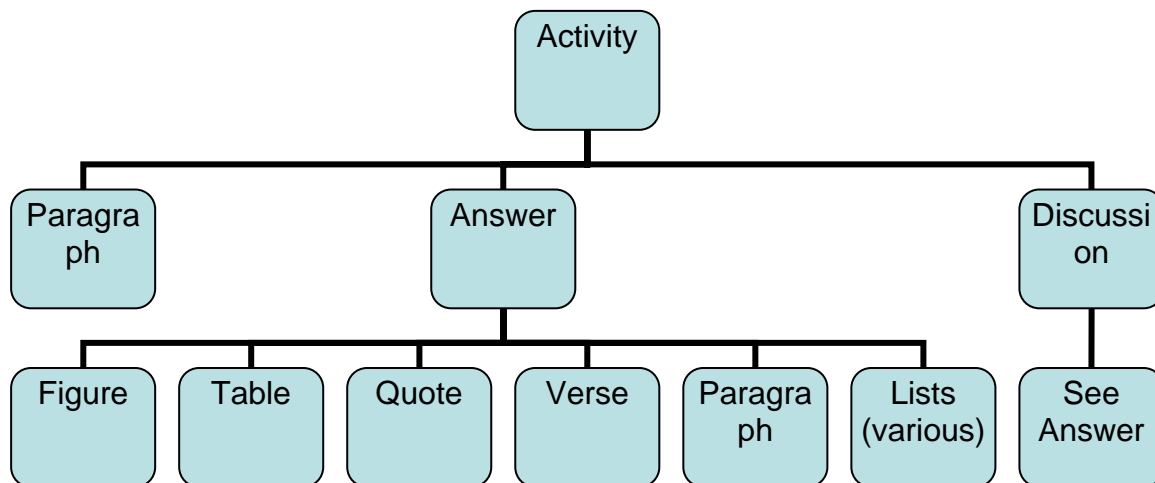




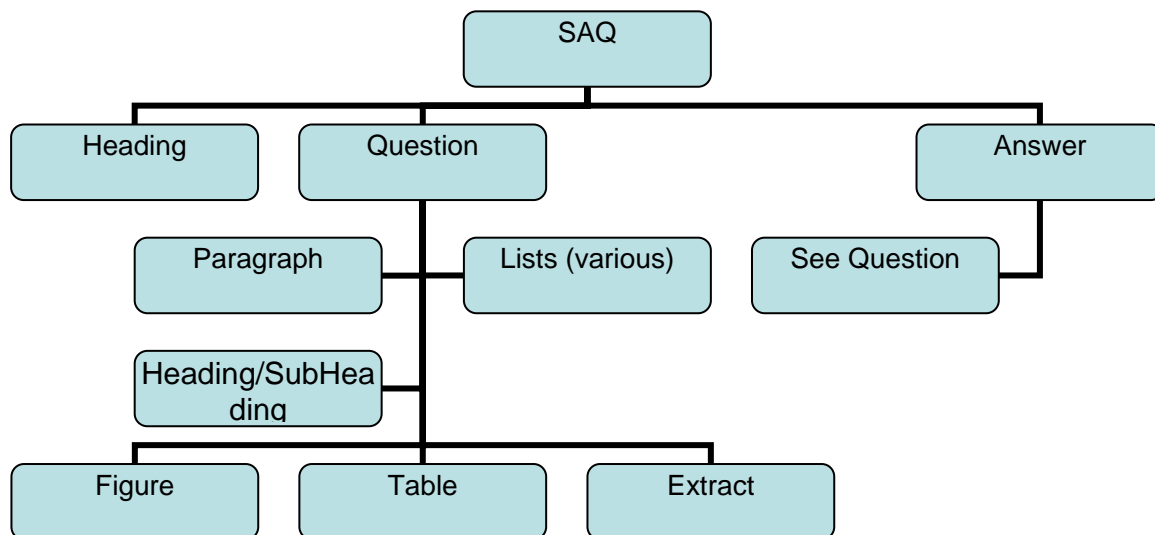
Subsection contains similar tags to that of Section.

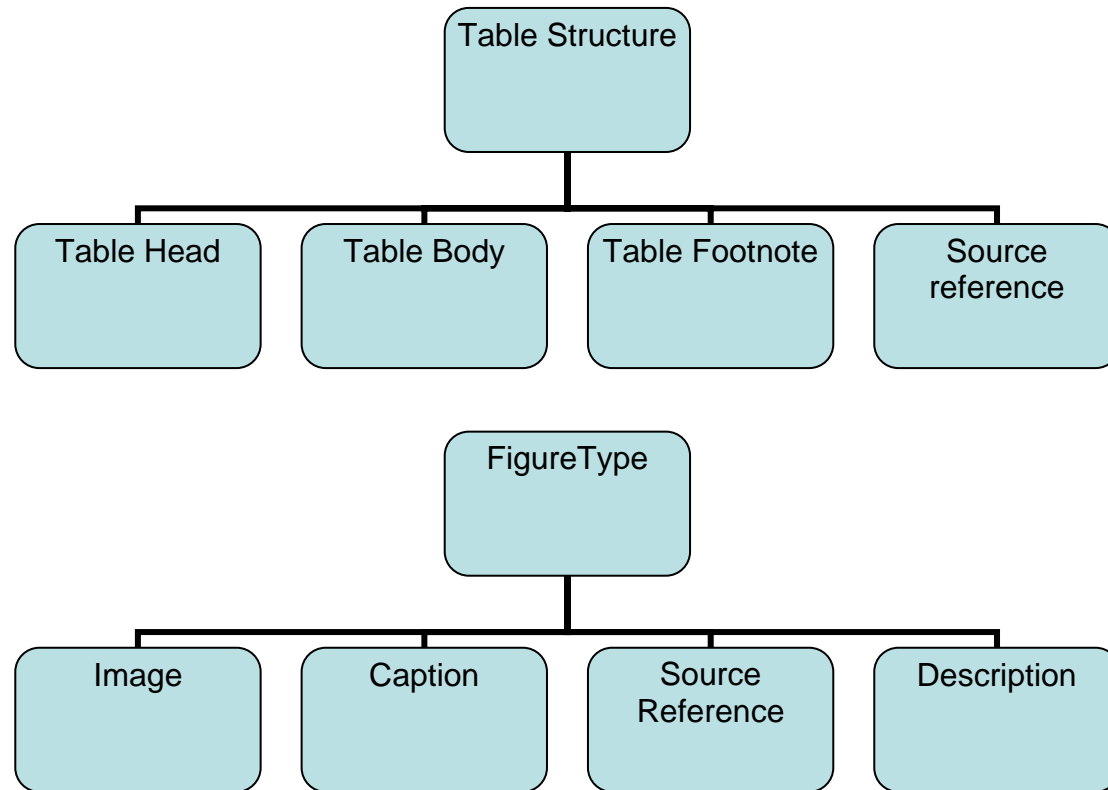


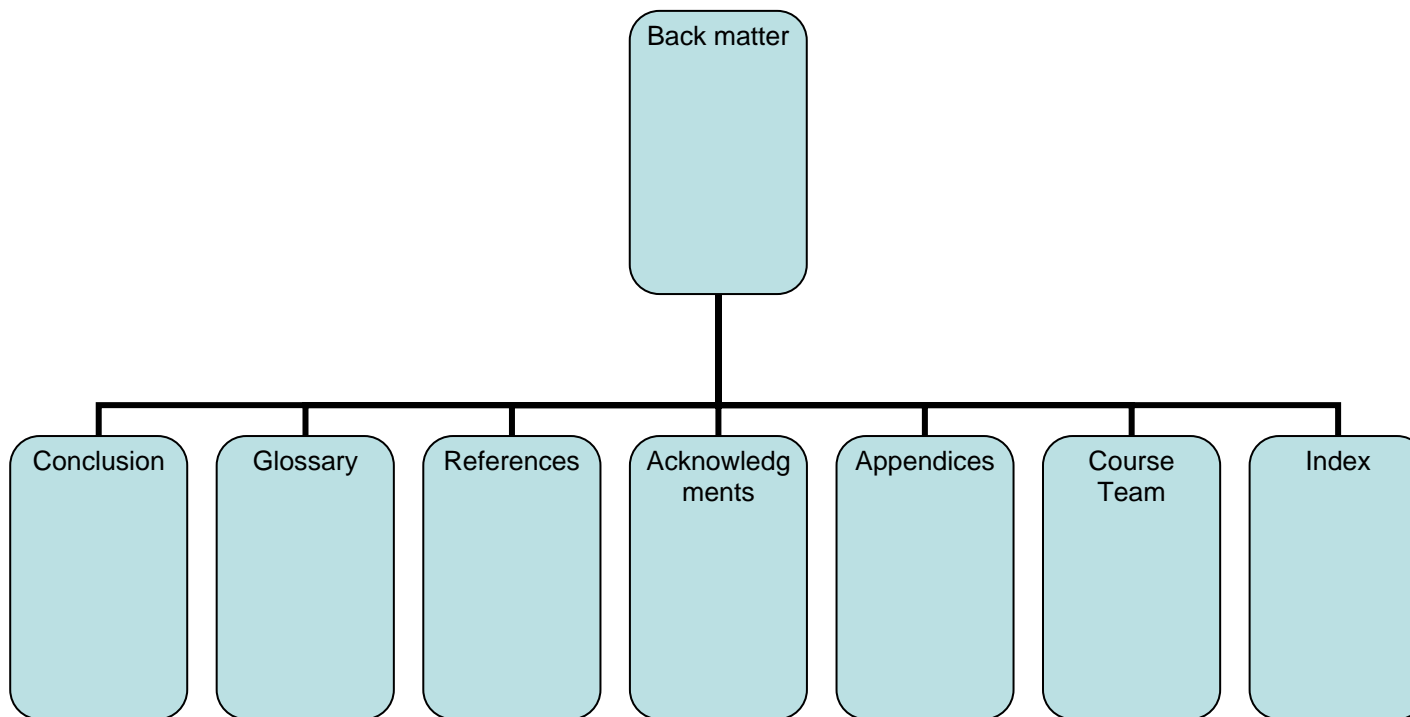
Some tags are very straightforward such as Heading and Subheading both of which may contain the b, i or u tags.



Exercise contains similar tags to Activity.











## Examples of tags in use

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<Item xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="OUGenericFull.xsd">
  <FrontMatter>
    <CourseCode>e.g. T173</CourseCode>
    <CourseTitle>full course title e.g. Technology: A Level 1 Course Engineering the Future</CourseTitle>
    <ItemID>block number the unit is taken from, and part or chapter of the course which comprises the unit</ItemID>
    <ItemTitle>originating book and block title</ItemTitle>
    <Imprint>
      <Standard>
        <GeneralInfo>
          <Paragraph>this section is information normally contained on an imprint page with space for any general
text</Paragraph>
        </GeneralInfo>
        <Address>
          <AddressLine>For example, The Open University, Walton Hall, Milton Keynes, MK7 6AA</AddressLine>
        </Address>
        <FirstPublished>Year</FirstPublished>
        <Copyright>&#xA9; [year as listed above] The Open University</Copyright>
        <Rights>
          <Paragraph>This will be standard rights information e.g.:All rights reserved. No part of this publication may be
reproduced, stored in a retrieval system, transmitted or utilized in any form or by any means, electronic, mechanical,
photocopying, recording or otherwise, without written permission from the publisher or a licence from the Copyright
Licensing Agency Ltd. Details of such licences (for reprographic reproduction) may be obtained from the Copyright
Licensing Agency Ltd of 90 Tottenham Court Road, London W1P 0LP.</Paragraph>
        </Rights>
        <Edited>the Open University</Edited>
        <Typeset>the Open University</Typeset>
        <Printed>Printed and bound in the United Kingdom by Burlington Press (Cambridge) Ltd.</Printed>
        <ISBN>0 7492 8454 4</ISBN>
        <Edition>Second edition</Edition>
      </Standard>
    </Imprint>
  </FrontMatter>
```



```
<Unit>
  <UnitID>unit unique identifying code, i.e., K100_1</UnitID>
  <UnitTitle>Title of unit</UnitTitle>
  <Session id="SES001">
    <Title>The entire unit should sit within session tags - they mark the session of learning. Title of unit
    (repeated)</Title>
    <Section id="SEC001">
      <Title>Each section should sit within section tags. Section title</Title>
      <Paragraph>Main text will appear here. Each new paragraph of text should have new paragraph
      tags.</Paragraph>
      <Paragraph>This would be a new paragraph.<i>This sentence would appear in italics and should always be
      within a paragraph</i>
      <b>This sentence would appear in bold and should always be within a paragraph</b>
      <u>This sentence would appear underlined and should always be within a paragraph</u></Paragraph>
    </Section>
  </Session>
```

```
<Section id="SEC002">
  <Title>New sections should always have a title</Title>
```

```
<Paragraph>You can use numbered, unnumbered or bulleted lists. This would show that the list items should be
bulleted</Paragraph>
  <BulletedList>
    <ListItem>
      <Paragraph>This would be a list item</Paragraph>
    </ListItem>
    <ListItem>
      <Paragraph>This would be another list item</Paragraph>
    </ListItem>
  </BulletedList>
  <Paragraph>You can also use numbered lists</Paragraph>
  <NumberedList class="alpha">
    <ListItem>
      <Paragraph>This would appear as item 1</Paragraph>
    </ListItem>
    <ListItem>
      <Paragraph>This would appear as Item 2</Paragraph>
    </ListItem>
  </NumberedList>
```



```
<UnNumberedList>
<ListItem>
<Paragraph>This would create an unnumbered, unbulleted list item</Paragraph>
</ListItem>
</UnNumberedList>
```

```
<SAQ id="SAQ001">
<Heading>
<b>SAQ 1.1</b> &#x0028;Learning outcome 1.1&#x0029;</Heading>
<Question>
<Paragraph>An SAQ is a self-assesment question. SAQ question appears here</Paragraph>
</Question>
<Answer>
<Paragraph>SAQ answer text</Paragraph>
</Answer>
</SAQ>
```

```
<SAQ id="SAQ001">
<Heading>
<b>SAQ 1.1</b> &#x0028;Learning outcome 1.1&#x0029;</Heading>
<Question>
<Paragraph>An SAQ is a self-assesment question. SAQ question appears here</Paragraph>
</Question>
<Answer>
<Paragraph>SAQ answer text</Paragraph>
</Answer>
</SAQ>
```

```
<Activity id="ACT001">
<Question>
<SubHeading>Question 1.1</SubHeading>
<Paragraph>Activities will have a similar structure</Paragraph>
</Question>
</Activity>

</Section>

<Section id="SEC003">
<Title>Section title</Title>
<Paragraph>Main text including reference to a figure. Reference to figure is preceded by<a href="FIG001"
type="figure"/>Figure 1.1. These tags contain details of where the image Jpeg files are stored.</Paragraph>
```



```
<Figure id="FIG002"><Image id="FIG002" src="FIG002.jpg"/>
<Caption>
<b>Figure 1.1</b></b>All figure details are contained between figure tags. Image tags contain references to image storage
location. Figure numbers are in bold. The description should be contained here.</Caption>
<Description>A description may be added for users with disabilities</Description>
</Figure>
<Box id="BOX001">
<Heading>Title of material in box</Heading>
<Paragraph>Text that appears within box tags will be boxed. Main text for box. Figures and lists work in the same
way</Paragraph>
</Box>
<Paragraph>
<FootNote>
<Paragraph>
<sup>1</sup></sup>Footnotes will appear in these tags but should be integrated into the text and the tags
deleted</Paragraph>
</FootNote>
</Paragraph>
<Paragraph>
```

```
<InlineEquation>
<Image src="IEQN001.jpg" width="col_100"/>
<Description>DESCRIPTION REQUIRED</Description>
</InlineEquation>
</Paragraph>

<Equation id="EQN001">
<Image src="EQN001.jpg" width="col_100"/>
<Description>1.1</Description>
</Equation>
```



```
<Table id="TAB006">

<TableHead>
<b>Table 1.6</b> the table heading appears here</TableHead>
<TableBody>
<table>
<tr>
<th>table heading</th>
<th>
</th>

<th>Example</th>
</tr>
<tr>
<td>data</td>
<td/><td/>
</tr>
</table>
</TableBody>
<SourceReference/>
</Table>

</Section>
```

```
<Table id="TAB006">

<TableHead>
<b>Table 1.6</b> the table heading appears here</TableHead>
<TableBody>
<table>
<tr>
<th>table heading</th>
<th>
</th>

<th>Example</th>
</tr>
<tr>
<td>data</td>
<td/><td/>
<Section id="SEC004">
```

```

<Title>Section title</Title>
<Paragraph>This would be a new paragraph.</Paragraph>
</Section>

<Figure id="FIG003"><Image id="FIG003" src="FIG003.jpg"/>
<Caption>A title, short explanation, or description accompanying an illustration or a photograph</Caption>
<SourceReference></SourceReference>
<Description>A textual description of the image</Description>
</Figure>
</Session>

<BackMatter>
<Conclusion><Title>DESCRIPTION REQUIRED</Title>
<Paragraph>Any conclusion should be included here.</Paragraph>
</Conclusion>
<Glossary><GlossaryItem>
<Term>Any glossary terms should be added here. Note these are not picked up in the main site Glossary</Term>
<Definition>DESCRIPTION REQUIRED</Definition>
</GlossaryItem>
</Glossary>
<References><Reference>Further reading lists should be displayed in these tags following house style rules:
Author(s) (date)<i>Title in italics</i>publisher</Reference>
</References>
<Acknowledgements>
<Paragraph>Any acknowledgements should be added here</Paragraph>
</Acknowledgements>
<CourseTeam>
<Paragraph>Course team information</Paragraph>
</CourseTeam>
<Index><IndexTerm>In the unlikely event that an index is included, terms should appear like this</IndexTerm>
<SubTerm>DESCRIPTION REQUIRED</SubTerm>
<SubSubTerm>DESCRIPTION REQUIRED</SubSubTerm>
</Index>
</BackMatter>
  </Unit>
</Item>

```