MONASH University
Information Technology

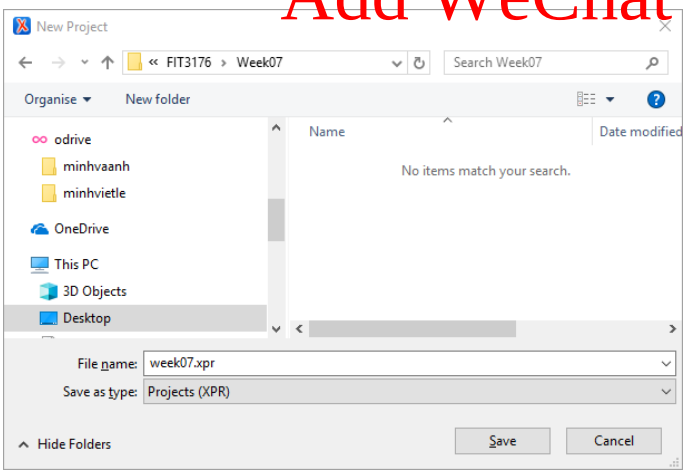
FIT3176 Advanced Database Design

Week 7 Studio - XML Documents

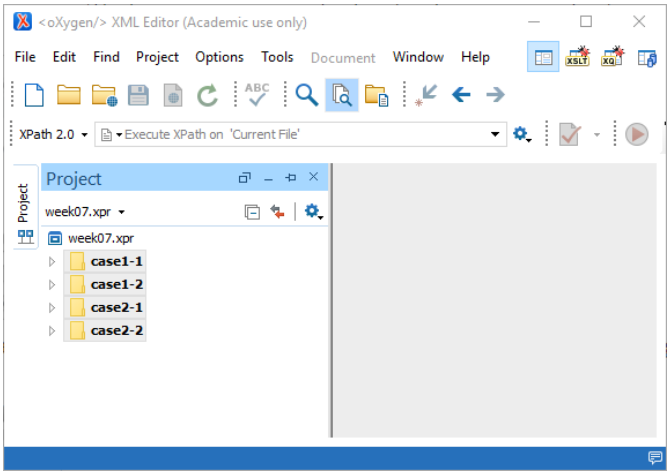
Throughout this section of FIT3176 we will be making use of oXygen - if you wish you can install this software at home, the software has already been installed in the on-campus labs. oXygen has been installed in the Mac labs and the Windows labs. If you install the software on your own computer please leave the defaults as you install, do not make oXygen the default for xml documents. oXygen software and its licence key can be downloaded from the Unit Information section in Moodle.

In the on-campus labs, you should set your workspace to a folder within your FIT3176 folder which you have been using for SQL Developer.

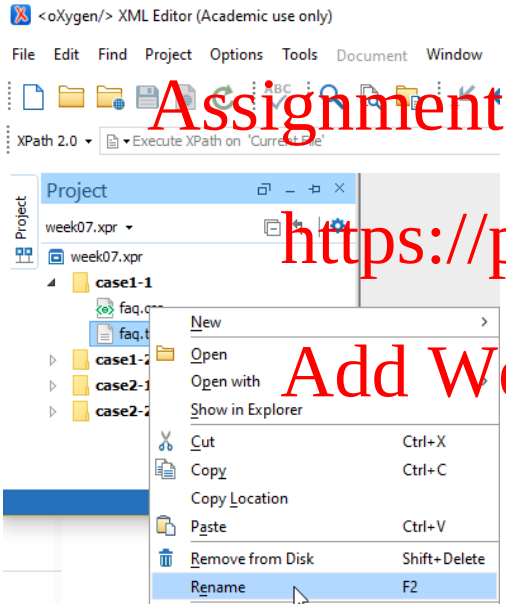
It is recommended that you create a new project in oXygen for each week's studio activities. Start oXygen and create a new project (Project - New Project). Locate this project in the week's folder (e.g. week7) and name it appropriately e.g. week7:



After the project has been created, you can drag data folders in your project folder from the file manager and drop them into the project or add them from the file manager to the project folder.



Each added folder can be expanded to show the files which are present in the folder/s (as shown above), files can be renamed by right-clicking the desired file:



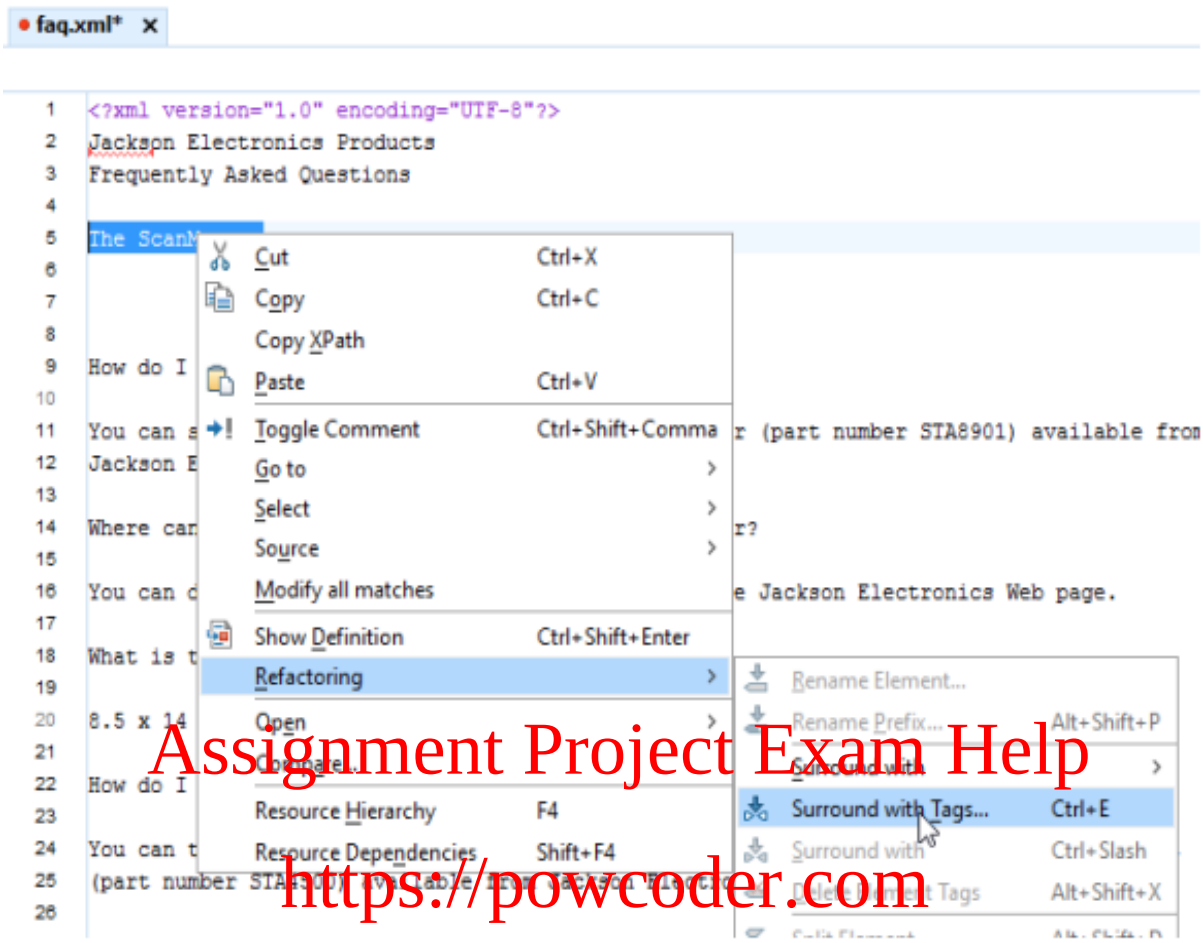
Some tips for using oXygen

(i) Adding a tag around a section of text:

● Highlight the text to be contained in the tags

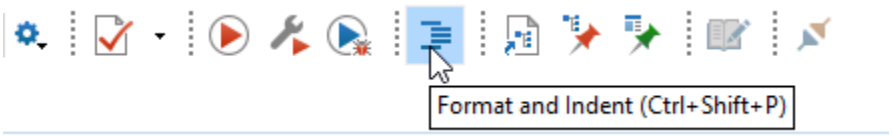
● Right click, select Refactoring - Surround with Tag… (or use the appropriate system shortcut, on the Mac Command-E, on Windows Ctrl-E)

● Enter the required tag

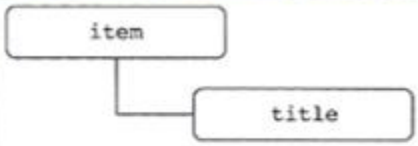
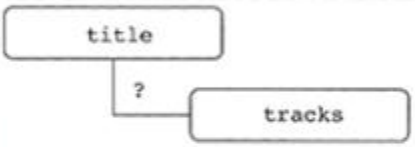
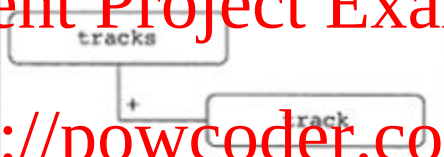


Note in this screenshot the file (faq.xml) has not been saved - this is indicated by the * alongside the file name. Many of oXygen's operations work on saved files, rather than what you have open in your browser - for this reason, be sure to regularly save your files.

(ii) automatically formatting the output: during data/tag entry do not be concerned with formatting the xml document. After all the tags/data are in place select the "Format and Indent" button in the toolbar:



Graphical Representation of XML documents

Symbol	Description	Chart	Interpretation
[none]	The parent contains a single occurrence of the child element.		An item can only have one title
?	The child element occurs once or not at all.		A title may or may not have a collection of tracks
*	The child element occurs any number of times.		The items element can contain zero or more item elements
+	The child element occurs at least once.		The tracks collection must contain at least one music track

Lab Activities

Reference: Carey, P., New Perspectives XML 2nd Edition Comprehensive, 2007, Tutorial 1 & 2, Selected Case Problems

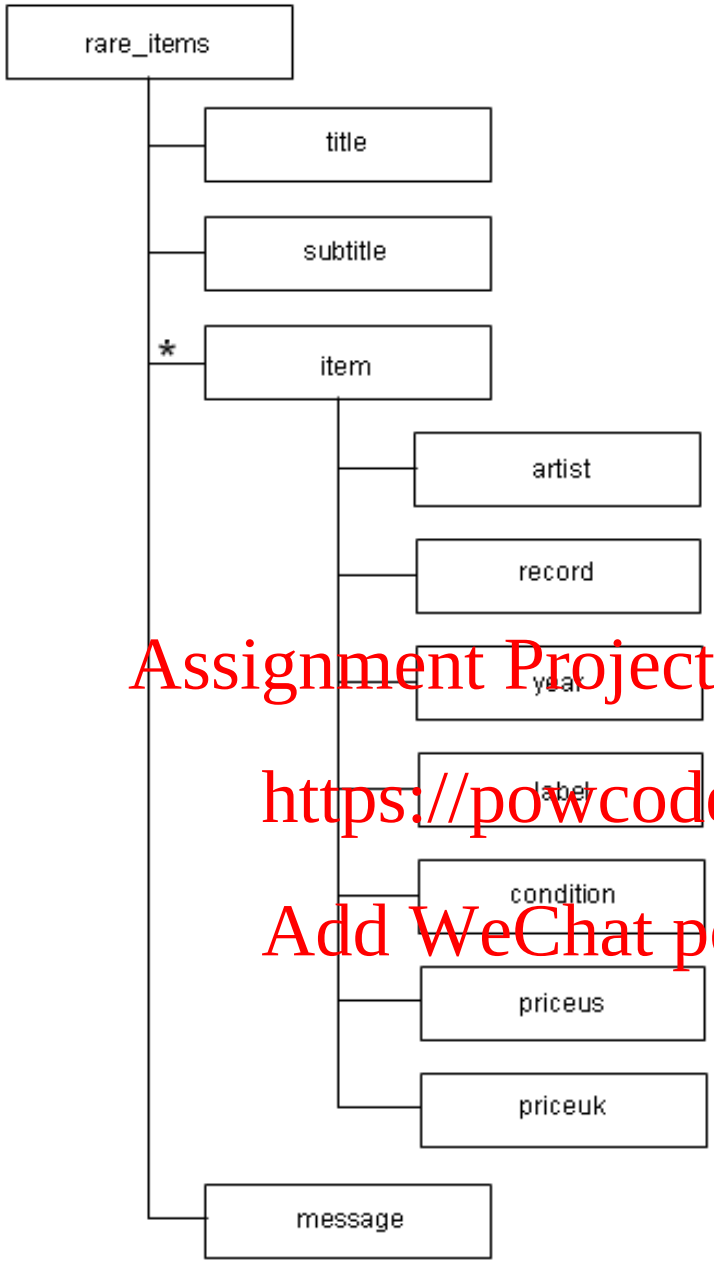
Several excellent web resources exist for checking character references, including

- https://dev.w3.org/html5/html-author/charref

Data files needed for the cases presented here are available in the week 7 data files archive available from Moodle.

Please note for FIT3176 you are not required to author CSS files, merely attach and use them.

Sample XML document



Examine the jazz and namespace examples discussed in the lecture and presented in Carey 2nd Edition Tutorial 1. Import and open the XML and CSS files in oXygen.

Examine the structure of the XML documents and look at the XML documents using your chosen external browser (this can be configured in the oXygen preferences).

Case Problem 1-1 Jackson Electronics

Jackson Electronics located in Santa Fe, NM, Jackson Electronics is a privately held manufacturer of consumer digital products such as scanners, printers, and digital cameras. Originally founded by Pete Jackson in 1948 as an office supply store, Jackson Electronics has thrived over the years with innovative thinking and effective use of cutting-edge technology. Alison Greely is one of the Webmasters for the Jackson Electronics Web Site. Her primary responsibility is to maintain information on the frequently asked questions (FAQs) section of the site. Alison would like to convert her documents into XML format and has asked for your help. She has given you a text file containing FAQs for two of Jackson Electronics' products: the ScanMaster scanner and the DigiCam digital camera.

Figure 1-25 below shows the structure she would like to apply to the text in her file.

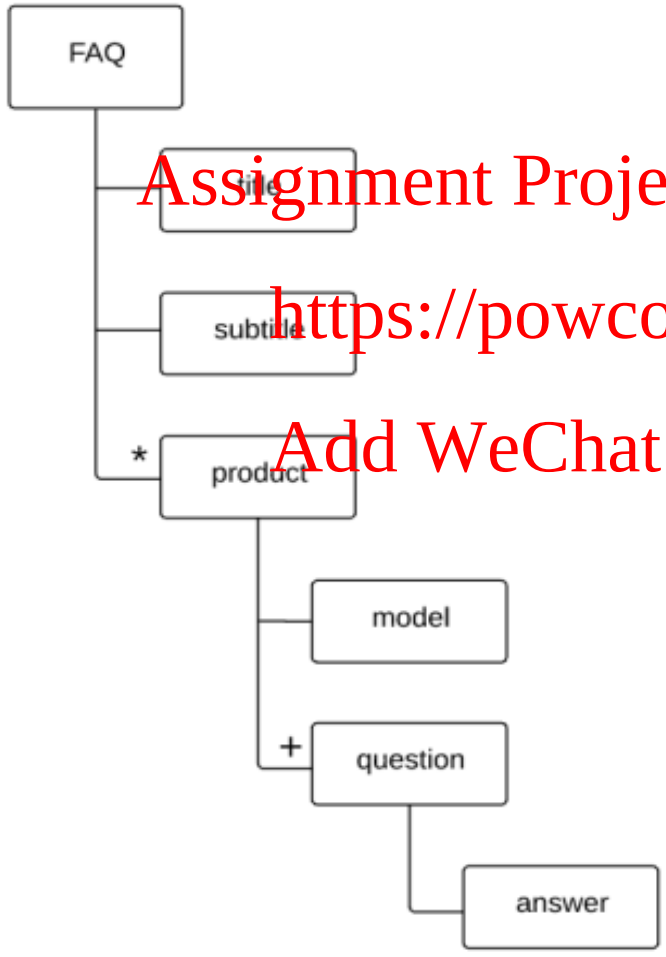


Figure 1-25 The structure for Case Problem 1-1 Jackson Electronics

She would like this text file converted to an XML document and then linked to a cascading style sheet.

To complete this task:

1. Import the case1-1 folder into your topic07 project and rename the faq.txt file to faq.xml
2. Create a prolog at the top of the document, indicating that this is an XML document using version 1.0 of XML (utf-8 encoding and standalone).
3. After the XML declaration, insert the following comment: “ScanMaster and DigiCam FAQ”. Add two more comments containing your name and the date.
4. Create a root element named FAQ.
5. Enclose the text “Jackson Electronics Products” in a title element, and set the text “Frequently Asked Questions” in an element named subtitle.
6. Create two product elements. Within each product element, insert a model element containing the name of the product.
7. Within the product elements, enclose each question in the text file within a question element.
8. Also within the product elements, enclose each answer in the document within an answer element, and place the text of those answers within a CDATA section.
9. Add a processing instruction to the prolog to apply the faq.css style sheet to this document.

Check your document is well formed using oXygen as you go, checking the final document in the external web browser. The final document should look like:

Jackson Electronics Products

Frequently Asked Questions

The ScanMaster

- How do I scan slides?
  You can scan slides using the JE Transparency Adapter (part number STA8901) available from Jackson Electronics.
- Where can I find the latest drivers for my ScanMaster?
  You can download the latest software drivers from the Jackson Electronics Web page.
- What is the largest sheet that I can scan?
  8.5 x 14 inches (216 x 356 mm).
- How do I fax with my ScanMaster?
  You can turn your scanner into a fax machine by purchasing the FaxRight add-on component (part number STA4500) available from Jackson Electronics.

DigiCam

- What is the difference between optical and digital zoom?
  With optical zoom, the Digicam's lens physically moves inside the camera; with digital zoom, the camera's processor zooms the image electronically. If you zoom too much, your image will become pixelated.

- What sort of batteries should I use with the DigiCam?
  Nickel Metal Hydride (NiMH) batteries work the best and having the longest lifetime. Nickel-Cadium batteries also work very well as do Alkaline batteries. Do not use Lithium batteries.
- What resolution should I use for 4x6-inch photos?
  We recommend 640x480 for 4x6-inch images, 1024x768 for 5x7-inch photos, and 1600x1200 for 8x10-inch photos.
- Can the DigiCam be harmed by airport X-ray machines?
  No, there is no evidence that X-ray machines can affect the performance or quality of DigiCam photos.

Case Problem 1-2 Midwest University

One of the original Federal Land Grant Universities, Midwest University now includes several world-class undergraduate and graduate programs. Professor David Teagarden is a member of the award-winning English department at MU. He is currently working on a Web site devoted to the work and life of William Shakespeare. He has created a document detailing the acts and scenes of Hamlet and has asked for your help in putting this data in XML format. Figure 1-27 below shows the tree structure that you'll apply to this XML document.

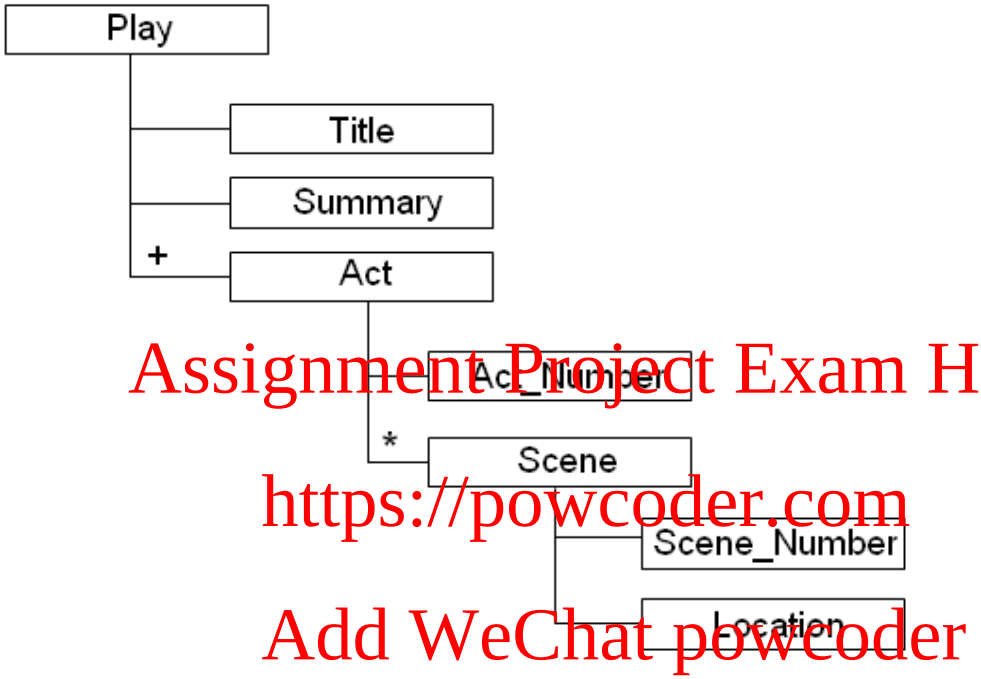


Figure 1-27 The tree structure for Case Problem 1-2 Midwest University

A style sheet has also been provided for you to display the Hamlet summary on the Web.

To complete this task:

1. Import the case1-2 folder into your topic07 project and rename the hamlet.txt file to hamlet.xml
2. Create a prolog at the top of the document, indicating that this is an XML document using version 1.0 of XML.
3. Add comments that include your name and the current date.
4. Enclose the entire document content in a root element named Play.
5. Place the title of the play in an element named Title. The Title element should be the first child of the Play element.
6. Add an attribute to the Title element named "type". Set the value of the type attribute to

"Tragedy".

7. Place the summary of the play in a CDATA section within a Summary element.
8. Place all of the information about each act of the play within an Act element. Place the name of each act (Act 1, Act 2, and so forth) within an Act_Number element.
9. Place all of the information about each scene of the play within a Scene element nested within each Act element. Place the name of each scene (Scene i, Scene ii, and so forth) within a Scene_Number element. Place the location of each scene within a Location element.
10. Create a processing instruction to attach the plays.css style sheet to this document.

Check your document is well formed using oXygen as you go, checking the final document in the external web browser. The final document should look like:

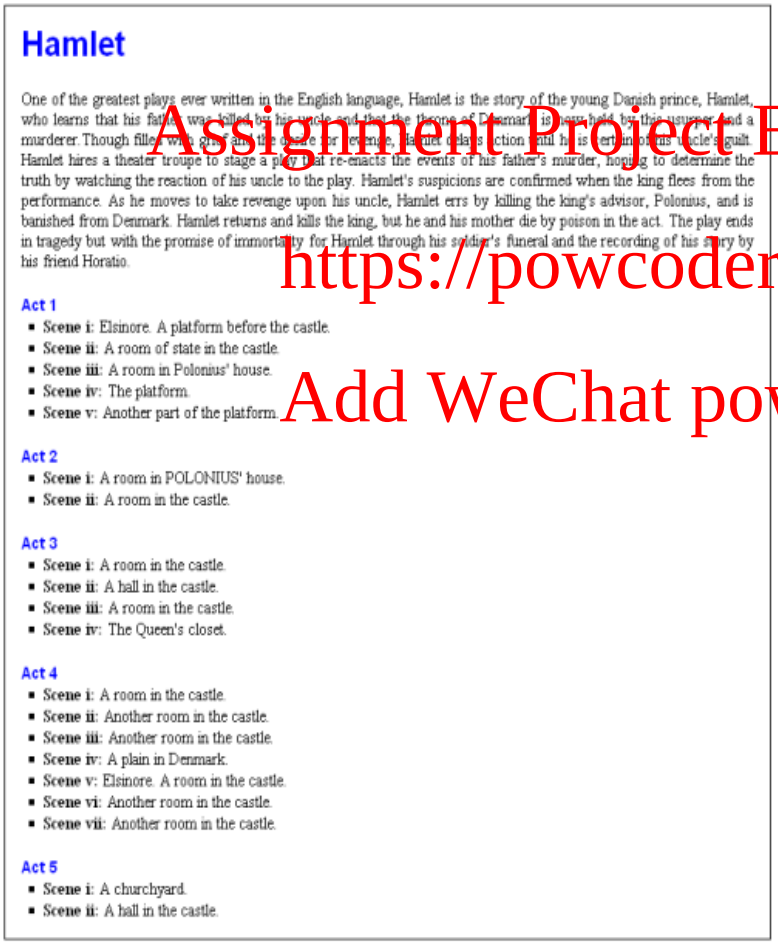


Figure 1-28 The completed Web page for Case Problem 1-2 Midwest University

Case Problem 2-1 EPAC-MO

EPAC-MO is an environmental political action committee located in central Missouri.The committee is currently in the midst of a fundraising effort that involves several canvassers travelling to the nearby towns of Cutler and Davidton to collect donations. Sudha Bhatia is managing the reports for the committee's intranet. She has placed some of the information in XML documents. She's developed one XML vocabulary that lists the employees or agents of the organisation that collect donations, and another XML vocabulary for the actual donations. She would like to create a compound document that lists the donations by each canvasser.

Figure 2-30 shows the structure of the compound document you'll create for Sudha.

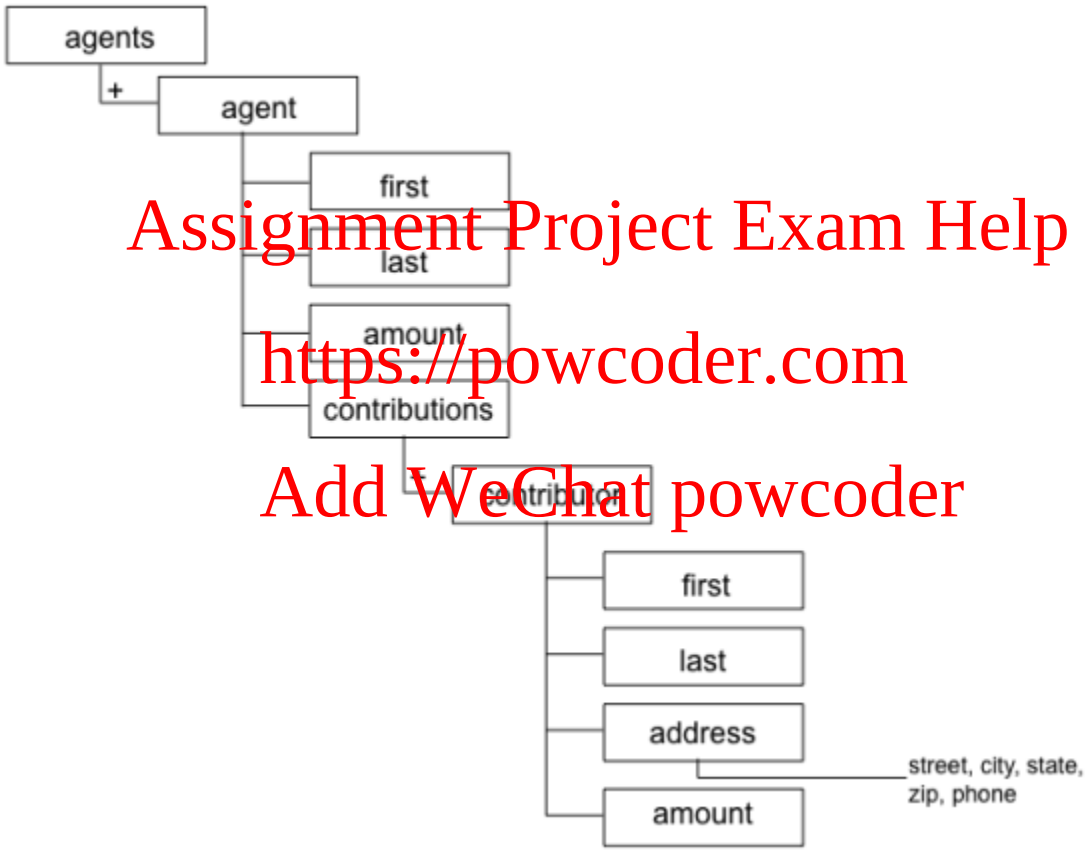


Figure 2-30 Case Problem 2-1

She wants the compound document placed within a Web page with different style sheets applied to the different vocabularies used in the page.

Figure 2-31 below shows a preview of the completed Web page:

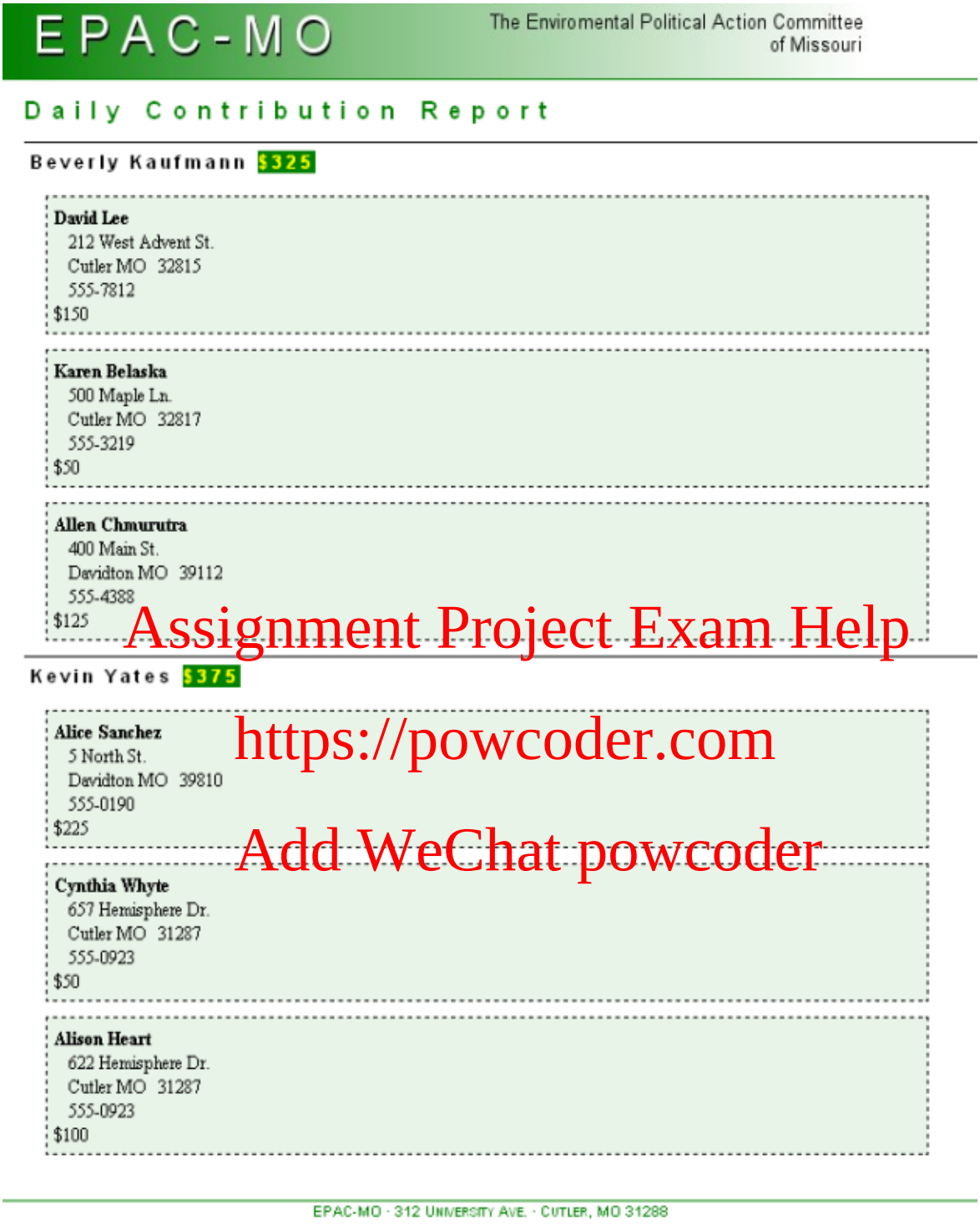


Figure 2-31 The completed Web page for Case Problem 2-1

To complete this task:

1. Open the report.html file with your text editor. Enter your name and the date in the comment section. This document is an XHTML document (this is an XML document with a HTML extension to allow a browser to correctly render images via the HTML <img> tag)
2. The default namespace of the elements in the document has been set to the XHTML namespace - xmlns="http://www.w3.org/1999/xhtml". Add namespace declarations for the:
3. agents namespace - "http://epacm.org/agents" with a prefix of ag and the
4. contrib namespace "http://epacmo.org/contrib" with a prefix of cont
5. Copy the elements from the agents.xml file and paste them into the report.html file directly below the h1 heading. Apply the agents namespace to each of the elements you pasted.
6. [NB this is a modified version of the text step] Copy the first agent's contribution from the contrib.xml file and paste them into the document directly above the first closing </agent> tag. (Hint: The first agent's contributions are identified by the agent attribute value "a01". Use the tree structure displayed in Figure 2-30 as your guide.) Apply the contributor namespace to all of the newly pasted elements.
7. [NB this is a modified version of the text step] Copy the second agent's contribution from the contrib.xml file and paste them into the document directly above the second closing </agent> tag. Once again, apply the contributor namespace to all of the newly pasted elements.
8. Close the report.html file, saving your changes and then open the document in your Web browser. Verify that the appropriate styles are applied to the elements from the XHTML, agents, and contributor vocabularies.

Case Problem 2-2 MathWeb

W3C Math Home - http://www.w3.org/Math/

The MathML namespace standard URI is: http://www.w3.org/1998/Math/MathML

Professor Laureen Cole of Coastal University is creating a Web site call "MathWeb" to use for online tutorials on mathematical topics. Laureen has been reading about the XML vocabulary MathML and how it can be used to display mathematical equations and information. She's asked you to create a compound XML document containing elements from XHTML and MathML.

A preview of the page that you'll create is shown in Figure 2-32 below:

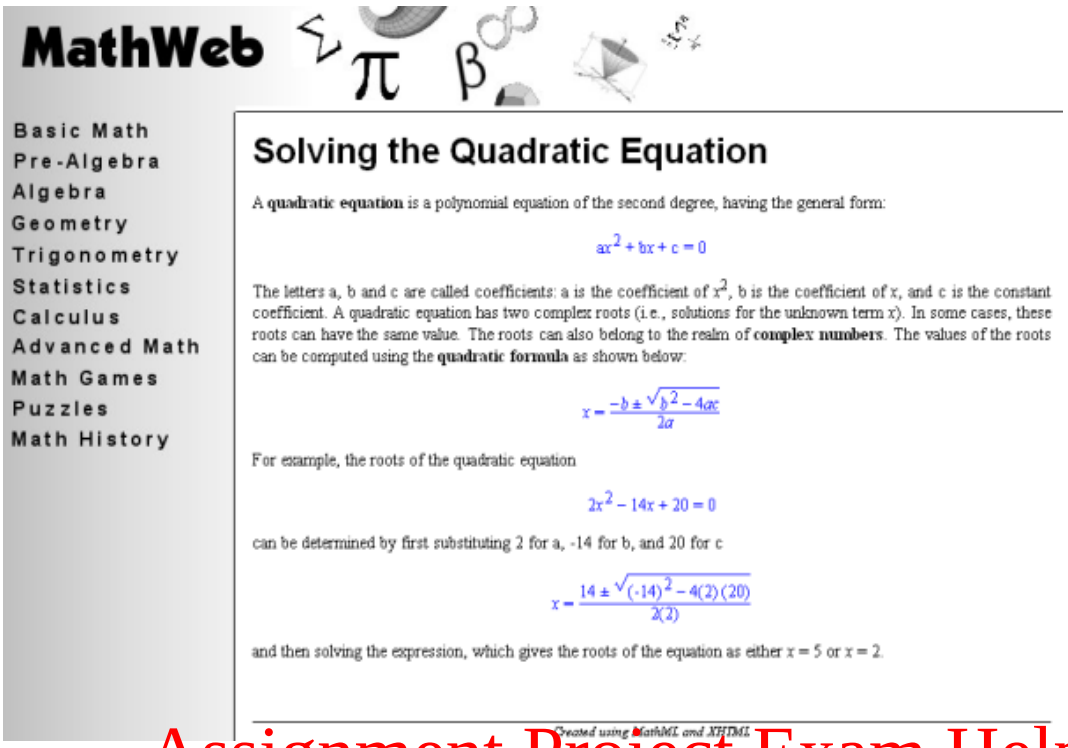


Figure 2-32 The completed Web page for Case Problem 2-2

Laureen has already created the content for the mathematical equations that she wants on her Web site. She wants your help in completing the Web page.

To complete this task:

1. Using your text editor, open quad.xml from the case 2-2 folder. Enter your name and the date in the comment section at the top of the file, and then save the document.
2. Within the root element, html, insert two namespace declarations for the XHTML and MathML namespaces. Make XHTML the default namespace of the document. Use the prefix "m" for all elements belonging to the MathML namespace.
3. Scroll down to the paragraph element with the id "eq1". Within this paragraph, paste the MathML elements from the eq1.xml file. Apply the MathML namespace to these pasted elements.
4. Repeat Step 4 for the paragraphs with the ids eq2 through eq4, pasting the elements from the eq2.xml through eq4.xml files. In each case, apply the MathML namespace to the pasted elements.
5. Close the file, saving your changes.
6. Open the quad.xml file in your Web browser. Currently, only Safari and Firefox desktop browsers support MathML natively (http://caniuse.com/mathml). Using one of these browsers verify that the MathML equations appear as shown in Figure 2-32.