




[newtFire {dhlds}](#)

Maintained by: Elisa E. Beshero-Bondar (ebb8 at pitt.edu)  Last modified: Sunday, 06-Aug-2017 18:16:42 EDT. [Powered by firebellies](#).

XPath Exercise 1

First of all, download the XML file I have linked here:

[ForsterGeorgComplete.xml](#). Open the file in oXygen and work with the XPath 2.0 Window. Respond to the XPath questions below in a text file, and [upload to Courseweb](#) for this assignment when you're finished. (Please use an attachment! If you paste your answer into the text box, CourseWeb may munch the angle brackets.) Some of these tasks are thought-provoking, and even difficult. If you get stuck, do the best you can, and if you can't get a working answer, give the answers you tried and explain where they failed to get the results you wanted. Sometimes doing that will help you figure out what's wrong, and even when it doesn't, it will help us identify the difficult moments. These tasks require the use of path expressions and predicates, and there may be more than one possible answer. You may opt to try the functions `count()` and `not()`, but they should not require any other XPath functions. Consult our introductory guide [Follow the XPath!](#) for help with constructing your expressions.

With the Georg Forster file open in oXygen and using the XPath 2.0 browser window in oXygen, construct XPath expressions that will do the following. **Be sure to give the FULL XPath expression you used in your answer, and don't just report your results.** This way, if the answer is incorrect, we can help explain what went wrong.

1. Like most of the long voyage publications, Georg Forster's voyage account is produced in multiple books, and inside each books we find multiple chapters. Both books and chapters are coded with `<div>` elements. Take a look at the outline view of the document before you begin to familiarize yourself with the structure of this file, and answer the following:
 - How can XPath tell apart the books from the chapters?
 - What XPath would find ONLY the books in the file?
 - What XPath would find ONLY the chapters in the file?
 - What XPath would find ONLY the chapters in Book 2?
2. Look at the outline structure of the document to help you with these: What's the XPath to identify the `<head>` element inside a chapter `<div>`? How would we locate a `<|>` (or line) element inside a chapter `<div>`?
3. Georg Forster used a lot of footnotes in his document: These are coded inside `<ref>` elements throughout the body paragraphs of the text. What's the XPath to locate all the notes in the document?

4. We've encoded lots of <placeName> elements in this document to mark names of places, and these may occur in lots of positions. Sometimes they're in the <head> elements that start the book or chapter divs, positioned inside lines of texts (coded with <l>). Most often they're nested in the body paragraphs (<p>), and they're frequently coded in Forsters notes, which you've just located.
- What's the XPath to determine the number of placeNames that appear inside ONLY the <head> elements, and not in the rest of the document? (Notice where these are located in the heads).
 - What's the XPath to find the placeNames that are only mentioned in the notes?
 - What's the XPath to find the placeNames that are only mentioned in the notes of Book I? How many of these are there?