|  |  |
| --- | --- |
| Aufgabe 2-0: BASE XML SETUP | |
| Steps Performed | Download the base XML for Team 13:   1. Original download: <http://online.univie.ac.at/vlvz?kapitel=707&semester=S2015&format=xml> is empty, since study programme changed, communication with Mag. Mangler and adaption to new URL: <http://online.univie.ac.at/vlvz?kapitel=706&semester=S2015&format=xml> 2. Cleaning of XML through search/replace for XML entities with the following results  Replace &amp;apos; -> &apos; 721 occurencesReplace &amp;quot; -> &quot; 15 occurencesReplace &amp;amp; -> &amp; 507 occurencesReplace &amp;lt; -> &lt; 0 occurencesReplace &amp;gt; -> &gt; 0 occurencesremove nbsp; in Line 466 -> This was a random space in the middle of a wordCorrected numeric entities &amp;# -> &# 439 3. Development of XSL based transformation in 3 steps to recognize the hierarchical structure and transform to English tagnames.   All 3 xsl transformations are performed using xsltproc XSL\_FILE INPUT\_XML > OUTPUT.XML  1 extractHierarchyAndRename.xsl: clean anchor levell at the beginning of file, rename tag names for numeric chapters to tag “chapter” and add a specific attribute to identify the level of a chapter xsltproc 1\ extractHierarchyAndRename.xsl baseXML\_706\_download.xml > output\_step1.xml  2 constructParentRelation.xsl: find the parent level ID through the order of chapters and add parentID attribute (preceding-sibling::chapter[@level = $parentLevel][1]/@id) xsltproc 2\ constructParentRelation.xsl output\_step1.xml > output\_step2.xml  3 constructHierarchyAndRename.xsl create the hierarchy as suggested and perform language transformation. Perform any kind of calculation and information structure needed for exercises. xsltproc 3\ constructHierarchyAndRename.xsl output\_step2.xml > output\_step3.xml  Pretty-print result XML xmllint --format output\_step3.xml > baseXML\_tranformed.xml |
| Solution | ../../../../../../../Desktop/Screen%20Shot%202015-04-04%20at%201 |
| Source Files | XML DOWNLOAD: <https://raw.githubusercontent.com/fridlutz/UNIVIE/master/INTERROP/Session%202/0%20XML_Clean/baseXML_706_download.xml>  XML RESULT: <https://raw.githubusercontent.com/fridlutz/UNIVIE/master/INTERROP/Session%202/0%20XML_Clean/baseXML_706_tranformed.xml>  DEV ENVIRONMENT: <https://github.com/fridlutz/UNIVIE/tree/master/INTERROP/Session%202/0%20XML_Clean> |