

XV UNL School, 21-25 July 2014, Geneva

Exercise #1 - Corpus

- Goal: To prepare the training corpus
- Deliverable: *corpus_org_<ID>.txt*¹
- Activities:
 1. Download the corpus available for your working language at www.unlweb.net/wiki/XV_UNL_School.
 2. Extract, from the corpus, the articles 1-5, and save them in one single plain text file (.txt) in UTF-8 format. Your resulting file should look like the following (in your working language):

<p>Article 1 All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.</p> <p>Article 2 Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty.</p> <p>Article 3 Everyone has the right to life, liberty and security of person.</p> <p>Article 4 No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms.</p> <p>Article 5 No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.</p>
--

3. Upload the file to the UTK (UNLWEB>UNLDEV>UTK>CORPUS>CORPUS>ADD)

¹ Where ID is the participant code provided at www.unlweb.net/wiki/XV_UNL_School, e.g., corpus_org_ASA.txt is the file to be provided by Ali Safari.