Text Modification Tool Read Me
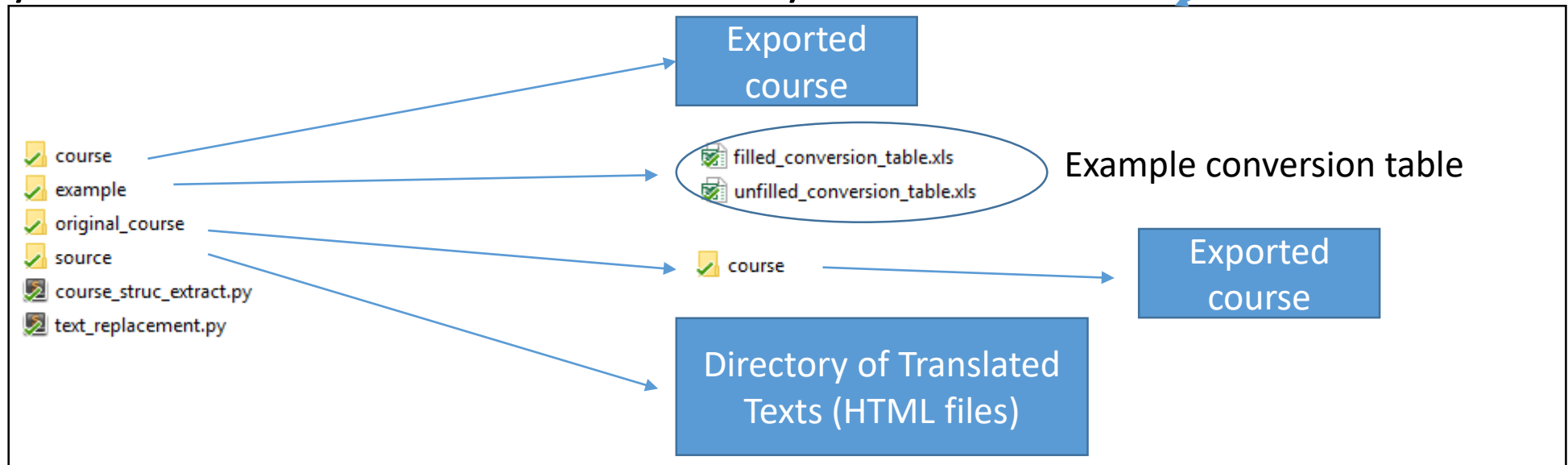
Workspace folder

- Contains 2 python scripts and example folder
  - course_struc_extract.py → this script generates an excel file 'conversion table'
  - text_replacement.py → this script replaces target text components to exported course
  - Example folder → contain an 'conversion table' example files both before and after filled a directory of translated text components

- Dependence python distribution
  - lxml, xlrd, xlwt, beautifulsoup4
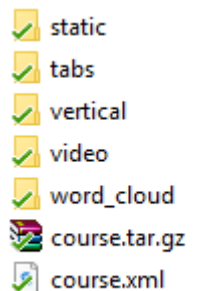
Files and Folders Preparation

- Unzip an exported edX course

- Create a 'course' folder
  - move exported course into this directory

- Create an 'original_course' folder
  - copy the 'course' folder into this directory

- Create a 'source' folder
  - Move all translated texts (HTML files) to this directories

Workspace folder should look like this

Exported course

filled_conversion_table.xls
unfilled_conversion_table.xls

Example conversion table

course

Exported course

Directory of Translated Texts (HTML files)

course
example
original_course
source
course_struc_extract.py
text_replacement.py

How to run the program

- Run a python script 'course_struc_extract.py'
  - excel file 'conversion_table.xlm' will be created

- Open the conversion_table and
  - manually add the path location of translated texts (HTML files) contained inside 'source' folder (please see example file for detail)

- Run a 'text_replacement.py'

- If program is successfully executed, <u>ready-to-upload course (course.tar.gz)</u> is created inside the course folder

static
tabs
vertical
video
word_cloud
course.tar.gz
course.xml

Notice

- This script is only tested with 1 edX course. Therefore, unexpected error should happen while using the program.

- There is a bug after uploading course back to edX. It is found that some text components disappeared in edX page. In our test, 1 component is missing.

- Some HTML component listed in 'error_files.txt' will not have figures showed properly in edX. This happens when original component (Eng) and translated component (local language) have different number of figures