ARLE LOMMEL, ALJOSCHA BURCHARDT



POST-EDITING AND ERROR ANNOTATION FOR PILOT 1 (SYNOPSIS)





WHAT YOU WILL SEE IN THE INTERFACE

Nr.	source	Question (x2)	Answer PE (x2)	Answer MQM (x2)
Segment number	The answer in English	Context (question) in your language	Answer for post-editing	Target for MQM annotation

When there are two columns for one item, the column with the pencil icon will be editable. **Even though the Question column is editable, you do** *not* **need to edit it.**





TARGET NUMBERS

Batch 2 QTLeap corpus	Target for Post-Edit	ing Target for MQM annotation
100 interactions for user evaluation Remaining	200 Post-Edits	100 MQM annot.
900 interactions		



GENERAL WORKFLOW

- First post-edit a segment and then do the error annotation (so that both are coherent)
- You can skip post-editing and annotation for trash segments
- You can do post-editing for a segment, but skip MQM if it seems unfeasible
- Start with the first 100 interactions and continue through the rest until you've reached the target numbers
 - There will be gaps, of course, when you skip.
 - Stop and talk to us when you feel that the whole task is overly time consuming; annotation speed differs a lot between annotators

4



FILTERING STRATEGY

- By default, you can annotate every segment and only skip completely unintelligible segments (or error annotations where it is simply pointless to try and identify individual errors)
- As Pilot 1 results are clearly no easy material, you may also apply filtering techniques. We do not want to impose any on the partners and leave it to you. If you filter segments, please document your criteria. Filters could be:
 - Annotate only "almost good" translations exhibiting 2-3 errors
 - Annotate certain linguistic phenomena that are of interest to you
 - Annotate segments with a limit in length
 - Etc.

