

General Guidelines/Questions for Translation Assessment

1. Does the translation retain the same grammatical structure and length of the original texts?
2. Does the translation lose any semantic information present in the original texts?
3. Are the names of places, people, items, events and dates retained in the translation?
4. Do the human-expert translations adhere to the same structure as the original texts in all the ways outlined in items (1,2,3)?
5. What could be improved in the machine translations and what could be improved in the human-expert translations?
6. Do you think the translation quality is sufficient to be considered for historical research?
7. Do you think the machine translations as they stand now (From the best performing models) are good enough to be trusted as utility
8. What are some questions or evaluation guidelines do you think are missing from my evaluation that could improve how the machine translations are assessed and evaluated?
9. What could be some useful metadata/extra information included in the translations that could be helpful when viewing the translations and utilizing them for historical research purposes?
10. How do you feel about using machine translations in the context of historical research
11. Do you think the original texts and human-expert translations hold in quality and were a good choice to use as a dataset for training and testing