

Research Institute in Information and Language Processing

Form 2

Ethics proposal form for (Level Six) Undergraduate written projects, Independent Study, Business and community link, and Master's dissertations.

2020/21

You should read the University Ethics Guidance before filling in this form. This is available from: <https://www.wlv.ac.uk/research/research-policies-procedures--guidelines/ethics-guidance/>

[For your information, the ACL currently adopts the ACM Code of Ethics and Professional Conduct: <https://www.acm.org/code-of-ethics>. This may be helpful when considering the ethical issues raised by your proposal].

Once completed fully, please submit this form to your supervisor for the project you are undertaking. Once the supervisor has approved and signed it, you must append a copy to your project placed as appendix A. Please note that this permission is only concerned with ethical issues and does not indicate anything about the intellectual merit of your project.

Please type details into the form.

1.	Name	[NAME]
2.	Student number	[NUMBER]
3.	Email address (this must be your University email address)	[EMAIL ADDRESS]

If this is a group project, please list ALL other students involved. (Full names and student numbers – number the new tables as 1.a., 2.a., 3.a.; 1.b.; 2.b, 3.b.; etc.):

4.	Subject to which the study will contribute:
Automatic Post-Editing, Translation Technologies	

5.	Name of supervisor(s)	[NAME]
----	-----------------------	--------

6.	Module code and title	7LN007/UM1, Dissertation (MA Computational Linguistics)
----	-----------------------	---

7.	Project title:
----	----------------

Research Institute in Information and Language Processing

Interactive Post-Editing via Online Adaptation of Automatic Post-Editing Models (*provisional title*)

8.	I confirm that I have: (Tick to confirm)	
	a.	Discussed my research with my supervisor. <input checked="" type="checkbox"/>
	b.	Read the Guide to Ethics and consulted the Ethics Guidance Web pages (https://www.wlv.ac.uk/research/research-policies-procedures--guidelines/ethics-guidance/) <input checked="" type="checkbox"/>

9.	Project research category: (Tick to as applicable)	
	Category 0	Research that does not involve human subjects or raise any ethical concerns. <input type="checkbox"/>
	Category A	Research that involves human subjects but is considered not to cause any physical or psychological harm. <input checked="" type="checkbox"/>
	Category B	Research that: <ul style="list-style-type: none"> • may be considered likely to cause physical or psychological harm. • may be contentious and/or risks bringing the University into disrepute. • requires accessing confidential data. • involves individuals considered to be vulnerable. Undergraduate and Taught Masters students are not normally permitted to undertake Category B research. <input type="checkbox"/>

Research Institute in Information and Language Processing

10.	The project involves: (Please tick all that apply)	
	Making video/DVD	<input type="checkbox"/>
	Making audio recording	<input type="checkbox"/>
	Observation of human subjects	<input type="checkbox"/>
	Participant observation	<input type="checkbox"/>
	Telephone and/or Email contact with individuals or organisations	<input type="checkbox"/>
	Interviews (structured/semi-structured/unstructured) [Delete as appropriate]	<input type="checkbox"/>
	Questionnaires (including on-line questionnaires)	<input type="checkbox"/>
	Access to confidential information	<input type="checkbox"/>
	Contact with minors (anyone under the age of 18)	<input type="checkbox"/>
	Contact with other vulnerable people (e.g. victims of crime, the recently bereaved)	<input type="checkbox"/>
	Research about a controversial issue	<input type="checkbox"/>
	Other [Please specify. E.g. Will project research outputs be open access? Will project research outputs be used to provide public services or services outside the University of Wolverhampton?]	<input checked="" type="checkbox"/>
Direct assessment (annotators are asked to assign scores to various translation options)		

11.	Brief outline of the project <i>[Expand as necessary. A quarter of a page may be sufficient.]</i>
<p>Despite the increasingly good quality of Machine Translation (MT) systems, machine-translated texts require corrections. Automatic Post-Editing (APE) models have been introduced to perform these corrections without human intervention. However, no system has been able to fully automate the Post-Editing (PE) process. Moreover, while numerous translation tools, such as Translation Memories, largely</p>	

Research Institute in Information and Language Processing

benefit from translators' input, Human-Computer Interaction remains rather limited in PE.

This research project discusses APE models and suggests that they could be improved in more interactive scenarios, as previously done in MT with the creation of Interactive MT systems. While some attempts were made to update APE models incrementally, this is often done using synthetic data, which is likely to affect the performance of the models. Therefore, in this project, APE models trained in a traditional (batch) setting were developed and updated in both batch and online modes without using any artificial resource, with the objectives of analysing the performance of online adaptation of APE systems. The experiments were conducted using outputs from different MT systems (statistical and neural), domains (generic and specialised), and language pairs (English-Spanish, German-English and English-Chinese). Developing an interactive interface to collect human post-edits was deemed out of the scope of this project, and the interaction with the translator was thus simulated using pre-existing references.

The results obtained with both batch and online models were analysed using automatic metrics (TER, BLEU, and other indicators such as precision) as well as via human evaluation (direct assessment).

12. Methodology

The description should also indicate:

- Your objective in gathering primary data from participants.
- How participants will be identified (including sampling method if doing questionnaires).
[For GDPR compliance, online surveys should be conducted using either JISC Online Surveys (JISC OS) or Microsoft Forms (via MS Teams). The University has site licenses for both.]
- The number of research participants.
- A sample of questions (if conducting either interviews or surveys).
- A sample of materials used to recruit research participants.

[Expand as necessary.]

The texts generated by all APE models (pretrained, updated in batch mode and updated in online mode) in all language combinations were extracted for human evaluation. Given the high number of segments in each case, a sample of 100 segments only was selected randomly for each model, thus accounting for 300 segments for each language pair. For comparison purposes, each segment is composed of a source sentence and three translation options: the MT output, the APE prediction, and the reference human post-edited version. To avoid bias, no information regarding the origin of these translations is shared with the participants. The evaluation procedure consists of assigning a score to each

Research Institute in Information and Language Processing

translation option, considering both adequacy and fluency, and commenting on possible recurrent error patterns. The primary objective of this analysis is to find out which variation of the APE models yields the best performance and to gain further insight on the results (compared to automatic metrics alone), in particular via the identification of recurrent error patterns.

The volunteers were recruited via emails sent by [SUPERVISOR] (see sample email below). They are fellow students of different levels (BA, MA, PhD) and backgrounds (e.g. linguistics, translation, computer science), with the exception of one participant, who is a business consultant. The workload is distributed differently across language pairs, with a total of 4 participants for English-Spanish, 3 for German-English and 6 for English-Chinese, thus accounting for a total number of 13 participants. In recognition for their collaboration, the participants are offered to have their names included in the Acknowledgements section of the dissertation and of a future publication as well as free attendance to events (co-) organised by RGCL members, such as the 'Technologies for Translation and Interpreting: Challenges and Latest Developments' seminar series (<https://em-tti.eu/em-tti-seminar-series/>) and the NeTTT conference (<https://nettt-conference.com/>).

It should be noted that the number of segments and the APE output assigned to participants is not specified in the dissertation. Any personal information in the annotated documents will be removed, so that these files can be open access.

While participants are asked to fill in a form, this is only to collect information about their background and whether they have any experience in translation. This information is summarised in the dissertation, but no detail allowing to identify participants is included.

- Sample email used to recruit participants:

[...]

May I ask you if you there are students who are speakers of Mandarin Chinese in August available at your university? I have the following request – a student of mine and I are working on a study/paper on automatic post-editing (APE) and the experiments include evaluation of the APE tool for German, Spanish and Chinese. We need native speakers of Mandarin Chinese to complete an evaluation task and were wondering if some of your students would be so kind to help with this.

The documents to be evaluated are a set of 3 Excel spreadsheets, containing 100 source (English) sentences each. Every source sentence comes with 3 translation options (in Mandarin Chinese). The task consists in assigning a score to each translation option, considering both adequacy and fluency, and commenting on possible recurrent error patterns. Therefore, there are 300 translation options in total to be evaluated for each file. There would be several possibilities for distributing the work, as one person could take 1, 2 or all spreadsheets, depending on everyone's availability. This task would need to be completed by the end of August.

The help of your students would be highly appreciated. In recognition for their collaboration, the names of the volunteers would appear in the Acknowledgements section of a future publication.

[...]

If you could simply forward this message to some of your students, I would be grateful.

Research Institute in Information and Language Processing

[...]

- Sample material to be evaluated by the participants:

Source	Translation A	Score (A)	Translation B	Score (B)	Translation C	Score (C)	Comments
But Moses threw the diadem down to the ground, and, in a puerile mood, he trod upon it.	但摩西把日记扔到地上，他又用一种卑鄙的心情踩着它。		但摩西把王冠扔到地上，很很一名了，离开它扔到了，很杀死到地上，很小扔到了，扔又没有去。		但摩西把王冠扔到地上，很孩子气地踩了上去。		
In c.1781 the facade was decorated with tufa by the Marchioness of Buckingham.	1781年，白金汉女骑士用圉法装饰了圉孔。		公元1781年，白金汉市的地方有凝灰岩。		公元1781年，白金汉市的侯爵夫人将外墙装饰有凝灰岩。		

13. Ethical Issues

Give a brief indication of the ethical issues (e.g anonymity, data protection etc) raised by your proposal and how you intend to address these issues.
[Expand as necessary.]

- Consent

Evaluators were asked to participate on a voluntary basis. While volunteers were offered certain opportunities (such as having their names in the Acknowledgements section of a future publication), they were free to decline the invitation to take part in this project.

- Data protection

Some personal details of the participants were collected in order to include information on the evaluation procedure in the dissertation (namely, background of the participants, their experience in translation and mother tongue). The names of volunteers will only appear in the Acknowledgements section and no information regarding the distribution of the work (e.g. who evaluated which file) is shared, as mentioned in section 12 of this document. Contact details of participants will not appear in the dissertation, nor in the open access documents.

- Workload

The participants were recruited between the 3rd and 25th of August 2021. The original deadline for this task was set to the 27th of August. The majority of volunteers (all volunteers for English-Spanish and German-English) responded by the 5th and were immediately provided with a file to evaluate, which was deemed to leave a sufficient period of time to complete the evaluation. As for English-Chinese, it took longer to find volunteers, which explains the higher number of participants for this language pair (the work was distributed among more volunteers to reduce individual effort). The volunteers who expressed their interest but were busy were offered to review a limited number of segments. All participants were told that they could get in touch if they had questions and request more time if needed.

Research Institute in Information and Language Processing

14.	Is ethical approval required by an external agency/parents of participants? <i>[If yes, please provide further details]</i>	<input type="checkbox"/>
	<div style="border: 1px solid black; padding: 5px;">No</div>	

15.	Is a DBS check required? If yes, please provide your disclosure letter together with this form. <i>[DBS checks are criminal records checks which can be requested online, for a fee (£23 - £40)). These are legally required for certain types of work involving vulnerable participants. If participants are recruited through a charity or through the NHS, these organisations may require you to undertake a DBS check. See https://www.gov.uk/request-copy-criminal-record for more details.]</i>	<input type="checkbox"/>
------------	---	--------------------------

Student Signature	
Date:	31.08.2021
Name of Supervisor	[SUPERVISOR]
Approval Signature of Supervisor	
Date:	31.8.2021