IMSISS Dissertation Feedback & Mark Sheet

Student Matriculation No.	**Glasgow** 2247793 **DCU** 17116155 **Charles** 22035308
Dissertation Title	Artificial Intelligence and Human Security: AI Strategy Analysis

INDIVIDUAL INSTITUTION GRADING

Reviewer 1 Initial Grade *For internal use only*	*Reviewer 2 Initial Grade* *For internal use only*	*Late Submission Penalty* *no penalty*
Word Count Penalty (1 UofG grade point per 500 words below/above the min/max word limit +/- 10%) **Word Count:** 21401 **Suggested Penalty:** *no penalty*		

JOINT GRADING (subject to agreement of the external examiner and approval at Joint Exam Board)

Final Agreed Mark. (Following correspondence reviewers should list the agreed final internal grade taking before and after any penalties to be applied).
Before Penalty: C3 [12]   **After Penalty:** C3 [12]

DISSERTATION FEEDBACK

Assessment Criteria	Rating
A. Structure and Development of Answer This refers to your organisational skills and ability to construct an argument in a coherent and original manner	
• *Originality of topic*	Good
• *Coherent set of research questions and/or hypothesis identified*	Good
• *Appropriate methodology and evidence of effective organisation of work*	Weak
• *Logically structured argument and flow of ideas reflecting research questions*	Good
• *Application of theory and/or concepts*	Good
B. Use of Source Material This refers to your skills to select and use relevant information and data in a correct manner	
• *Evidence of reading and review of published literature*	Good
• *Selection of relevant primary and/or secondary evidence to support argument*	Good
• *Critical analysis and evaluation of evidence*	Good
• *Accuracy of factual data*	Satisfactory
C. Academic Style This refers to your ability to write in a formal academic manner	
• *Appropriate formal and clear writing style*	Very Good
• *Accurate spelling, grammar and punctuation*	Very Good
• *Consistent and accurate referencing (including complete bibliography)*	Very Good
• *Is the dissertation free from plagiarism?*	Yes
• *Evidence of ethics approval included (if required based on methodology)*	Not required
• *Appropriate word count*	Yes

ADDITIONAL WRITTEN COMMENTS

Reviewer 1

The thesis "Artificial Intelligence and Human Security" is a good balanced analysis of the topic between qualitative and quantitative methods. The methodological aspects and discussion are definitely the most substantial parts of the thesis, which ground appropriately the conclusions. However, the main issue with the thesis is the difference between the first part, before the data collection and analysis, and the second part. The first could have been improved (some suggestions below) to be more consistence and to avoid minor misleading (subjective) statements. The evidence reported in the data analysis is straightforward and its analysis is appropriate. The critical analysis of the literature is sufficient to ground the main idea of Human Security, the discussion of it could have been more compact and shapr. Finally, the accuracy of the factual data is satisfactory.

Specific comments
1. The Turing's contribution was much wider than what is suggested in the introduction. Actually, he was the locician who invented the notion of computing machine itself (Turing Machine) and who implemented it in a real machine. In addition, he was also the developer of pionereeing studies on software etc.. The history of the Artificial intelligence is really too short to be relevant (and, it was, maybe, not so relevant in explaining the topic for the purposes of the thesis).
2. There are few strong statements that, though not completely inaccurate, could be misleading and they could have been referenced to avoid this potential point. These statemetns can be found in different parts of the thesis, which seems to suggest a trend in it (e.g. "For many decades AI remained primitive" p. 11, "Some of these changes have impact in trivial ways" p. 7, "This is understandable, as the field was in its adolescence, and the future seemed bright with possibilities" p. 11, "Though not a direct risk now at the current level of intelligence that AI has achieved, this is an emerging risk that we as a society need to prepare for as we push forward in development. A catastrophic existential threat is the most widespread possible impact on Human Security, potentially destroying us all", p. 75 etc).
3. The analysis of the Human Security (sometimes spelled in other ways) could have been improved: instead of giving just a list of items, a general discussion of them could have helped in making the point more clearly.
4. The thesis is linguistically well written, but some loose expressions (or slightly subjective statements) are still present (e.g. "Similarly, Khong feels that Human Security is a nice, idealistic notion" p. 15, "it does this by breaking down the high-level concept into more manageable (and more useful!) separate categories". P. 19 etc.).
5. Though the thesis is grounded on a consistent vision, it sometimes contains minor inconsistences (e.g. "Artificial Intelligence is complex and heavily technical, but well defined" p. 23 which seems to be in contrast with what is said at the beginning e.g ""Artificial Intelligence" is really an umbrella term for a set of technologies that attempt to act and make decisions autonomously" p. 8.
6. The methodological analysis and the discussion of the data and findings are quite substantial and consititue a strong base for the conclusions.

Reviewer 2

The dissertation offers a potentially interesting and valuable intersection between Human Security and Artificial Intelligence. Although the original idea is sound, the execution, and in some parts also the depth of the analysis, suffers substantial shortcomings. First, the paper's references to literature remain relatively sparse. This particularly concerns the theoretical and

methodological part outlining the concepts of AI, Human Security, as well as the applied methodology. If remedied, other issues could have been avoided. More precisely, the introduction of AI completely omits Machine Learning as a scholarly discipline together with its achievements. Further, and also based on the previous point, the proposed taxonomy mixes learning paradigms with the means of their realisation. Moreover, the paper seems to draw heavily, also mostly without references, on the scholarship founded by Bostrom, which is not in any way representative when considering the Machine Learning state-of-the-art. This is perhaps quite aptly illustrated by the following sentence "The emergence of AGI is inevitable on the path we are on" (p. 11). There is no sholarly consensus regarding the timeline, technological feasiblity and also the inevitability of the trajectory itself. Moreover, the sole concept of AGI is highly contested, since even the human intelligence couldn't be considered an example of general intellingece in its full breadth and depth. Most of the former general points apply also to the section on Human Security, which is much stronger, however, still a bit light on the literature making the part rather non-systematic. Finally, the methdology is not analytically convincing. Also, the meta-theoretical background (Dewey) lacks clarity, since it fleetingly mentions political realism instead of the philosophical notion of realism broadly construed that would be a natural counterpart for Dewey's debate. Overall, the dissertation suffers from various problems, however, it still elaborates an interesting link between AI and HI, which gets revealed in the emprical synthesis.