Part 1: Ideas from Medical Ethics for NLP Ethics

1.1: Rewriting 4 of the basic principles of the Declaration of Helsinki

- 13. The design and performance of each [NLP Application] involving [User(s)] should be clearly formulated in an [NLP Application] protocol. This protocol should be submitted for consideration, comment, guidance, and where appropriate, approval to a specially appointed ethical review committee, which must be independent of the [NLP Developer], the sponsor or any other kind of undue influence. This independent committee should be in conformity with the laws and regulations of the country in which the [NLP Application] is performed. The committee has the right to monitor ongoing trials. The [NLP Developer] has the obligation to provide monitoring information to the committee, especially any serious adverse events. The [NLP Developer] should also submit to the committee, for review, information regarding funding, sponsors, institutional affiliations, other potential conflicts of interest and incentives for [User(s)].

- 19. [NLP Application] is only justified if there is a reasonable likelihood that the [User(s)] in which the [NLP Application] is carried out stand to benefit from the results of the [NLP Application].

- 23. When obtaining informed consent for the [NLP Application] the [NLP Developer] should be particularly cautious if the [User(s)] is in a dependent relationship with the [NLP Developer] or may consent under duress. In that case the informed consent should be obtained by a well-informed [NLP Developer] who is not engaged in the investigation and who is completely independent of this relationship.

- 25. When a [User(s)] deemed legally incompetent, such as a minor child, is able to give assent to decisions about participation in [NLP Application], the [NLP Developer] must obtain that assent in addition to the consent of the legally authorized representative.

1.2: Examine, in detail, the appropriateness of each of your rewrites

- 13. This principle is still relatively applicable for NLP. Especially within NLP research, it is wise to have a protocol which is submitted to an ethical review committee if users are involved. However, during development of the NLP application, if users are not involved, having an ethical review committee might be over the top (but there is no harm in having one). Perhaps a revision could involve statements on how having a review committee is not always necessary given some circumstances, and that a review committee may possibly be stifling towards development.

- 19. This principle is very applicable for NLP. An NLP application should only be developed when the users benefit from the application (as opposed to it being a detriment). Every single aspect of the principle is applicable.

- 23. This principle is applicable to the NLP application when there are active users involved. Given active users, the principle is entirely applicable in that when giving consent, the user should not be in a dependent relationship with the developers. As with all applications or experiments, the user should be able to provide informed consent without any external pressure, and that they should use the application at their own will.

- 25. This principle is very applicable for NLP. As with all terms and services of any application, users deemed legally incompetent should also have the consent of the legally authorized representative to be able to use the NLP application. There should be no revision to this principle.

1.3: Review the transferability of the overall principles

Overall, I feel like the majority of the principles can be transferred to NLP. They contain vital information to ensure the health and safety of the users, while maintaining the morality of the researchers (or NLP developers in our case). Given that the Declaration of Helsinki was written for medical researchers and physicians, sometimes the principles are not all that applicable, or may be overwhelming, or may not even cover some NLP ethics which should be addressed.

An example of a principle which may seem overly protective is principle 18. I would disagree that sometimes in order to make progress within NLP, it is acceptable to have risks and burdens (as long as it is not undue) which outweigh the importance of the objective. Given that the volunteers have provided informed consent about all possibilities of risks and burdens under their own will, I think it is okay especially within NLP development. NLP innovation is tough as it is, and given a wider view of the world, NLP development (no matter the objective) is typically used to further our technological development as opposed to helping or saving a life (but it can be used for that as well). Sometimes it is fine if the risks outweigh the importance of the objective, as long as it is not undue.

Perhaps one principle or aspect within NLP that the set of principles does not cover is that the use of language data may be easily taken from public sources such as Twitter, Wikipedia, or Reddit. Users may not necessarily have knowingly provided consent that their data would be used within NLP development, and it should be addressed. An additional aspect is that deploying a model may turn from observational to interventional, and this should be mentioned to the users before they provide informed consent to participate.