**A Search Engine for Discovery of Biomedical Challenges and Directions - Annotation Guidelines**

**Introduction**

We discuss the annotation of scientific challenges and directions mentioned in the Biomedical literature.

In particular, our goal, with your help, is to build a search engine that allows a user to find sentences related to scientific challenges and directions. A user will search for a term and will see relevant sentences in context.

In the following section we present detailed guidelines for marking the annotations manually in text. Please read the guidelines carefully and adhere to them during annotation.

**Definitions & Examples**

You will need to decide if a sentence mentions a “**challenge”**, a “**research direction”**, both, or none.

**Challenge definition -** A sentence that unambiguously mentions a problem, difficulty, flaw, limitation, failure, lack of clarity, or a gap in knowledge.

Notes:

* Simply mentioning a disease, virus or a symptom is NOT a challenge. There needs to be a mention of some clinical difficulty, undesired outcome or serious detrimental effect in the context of a disease to qualify as a challenge.
  + Use your judgement about the distinction between “trivial” mentions of diseases and serious effects. Implicit references to the latter qualifies as a challenge.
* If the same sentence also suggests that the challenge has already been fully solved, then do NOT count this as a problem sentence.
  + Use your judgment about this - there are many ways to write that a problem is considered to have a solution/ no longer be a problem.
  + If the problem has been mitigated or partially solved, then it still counts as a problem.

Examples:

* *Little is known regarding the association between IFN-β deficiency and occurrence of cell death.*
* *In this study MBL Etest results were difficult to reproduce and overestimated the presence of MBL in a low prevalence country like Norway.*
* *Measles infection is a highly contagious, air-borne, acute febrile illness caused by a paramyxovirus of the Morbillivirus genus [1] and still remains an unresolved global issue with considerable mortality and morbidity rates [2]*

Negative Examples:

* *Each task has three difficulty levels: easy, medium, and hard.*
  + *The sentence simply mentions a fact not a challenge.*
* *In contrast, the mice in the control group survived the challenge.*
  + *Even though the sentence includes the word ‘challenge’, it describes a challenge to the mice, not to the researchers or to society.*
* *No antiviral activity was observed in the ferrets.*
  + *Even though a lack of antiviral activity may be a problem, it may also be what is expected in an experiment. Thus, the sentence does not unambiguously mention a challenge.*

**Research Direction -** a sentence that unambiguously mentions suggestions or needs for further research, hypotheses, speculations, indications or hints that an issue is worthy of additional exploration.

Notes:

* A research direction needs to **“constructively/proactively”** indicate what should/could be researched, and not only refer to a lack of knowledge/certainty. You need to be able to look at the sentence and immediately see what the implied question/direction for further research is.
* “Negative” statements that only talk about a limitation/lack of knowledge without the “constructive/proactive” aspect indicating a direction -- are challenges, not directions.

Examples:

* *An open question in many of the site-specific frameshifting systems described in prokaryotes and eukaryotes is the degree to which the event is directional and reading frame specific.*
* *This finding suggests genetic factors might be potentially important to infraction treatments.*
* *We need to determine the role of IL-6 in future work.*

Negative Examples:

* *On stainless surfaces the SARS viral load decreased after 48 hours.*
  + *The sentence simply mentions a fact not a direction.*
* *But as William Osler reflected, "Medicine is a science of uncertainty and an art of probability." So even though there is much that is still uncertain, we need to act prudently in a timely manner rather than waiting for perfect information.*
  + *Even though the sentence includes an outlook on research there is no clear proactive direction that is mentioned.*
* *The number of animals who suffered is surprising.*
  + *Even though the result was unexpected it does not unambiguously indicate a clear research direction.*

**General Notes**

* **When in doubt for a given sentence,** ask yourself:*Assume I’m searching for challenges/directions related to a topic in this sentence. Would I be generally satisfied with this sentence being shown?*
* **Don’t spend too much time per sentence -** should be on average no more than 30 seconds (could often be less, or sometimes more if you read context or the sentence is particularly difficult to parse, etc.)
* **Regarding context - you can use the context if you are unsure about a sentence and you think having additional context is helpful**.
* If you read the context, add a “1” to the *looked at context* column
* If context sentences refer to challenges/directions, but the main sentence does not -- not a challenge!
* **External knowledge** - try to minimize using external knowledge as much as possible
* For one example, if you see a paper from 20 years ago mentioning a problem that you know was solved in 2020, do NOT use this knowledge.

**That's it! You can reach out if there are any questions.**