

Now that we're about halfway through the task I just wanted to send a reminder about assessing and some clarification about the task.

By now, you've all used the preference judging system. For some of you, the preference judging has involved a large number of judgments, and we're going to work to remedy this for the remaining topics such that for further topics **you will only do preference judging on documents that you have judged to be VERY-USEFUL and which contain a correct answer to the topic's question.** As a reminder, I'll review the difference between USEFUL vs. VERY-USEFUL.

The definition of USEFUL: The user would find the document useful because it either **directly answers the question or provides enough information for the user to determine an answer.** Some questions ask about the effectiveness of a specific treatment for a health issue, and merely mentioning the health issue or treatment of the question is not-useful. **To be useful, a document must address all of the parts of a question and help the user make a yes/no decision for the question.**

The definition of VERY-USEFUL: In addition to helping the user make a decision about the question's answer, **the document is high quality either because of the detail with which the question is addressed and/or the document appears to be from a highly credible source.** This document is something that you think **deserves to be in the top 10 results** of a web search for this topic's question.

While both useful and very-useful documents address the entire question and help the user determine an answer to the question, very-useful documents are of "top 10 web search" quality because of their answer quality and/or apparent credibility of their source. You can find more than 10 very-useful documents for a given topic, i.e. you are trying to determine candidates for being in the top 10 results.

When I say "credible", I mean does the document give you confidence in the answer it gives? It might come from an authoritative source, or cite authoritative references, or be written extremely well, or clearly not be selling something, or be written by someone who seems to be an expert. Remember, THIS IS NOT CORRECTNESS – you can have a credible document telling you that injecting bleach is good for you, and that document would cite a bunch of studies and be written so clearly and strongly that it inspires confidence in the answer.

After you have judged a document's usefulness, you must judge what the document says the answer to the topic's question is. All questions are posed as yes/no questions. You are recording whether the document says the answer is "yes" or says the answer is "no". For example, if the question is "Does yoga help arthritis?", a "yes" document tells the reader to do yoga because it will help their arthritis, and a "no" document would tell the reader to not bother with yoga because it doesn't help arthritis. Labeling the documents with "yes" or "no" is very important, for you will only be shown the VERY-USEFUL documents that contain a correct answer to the question when you do preference judging.

At the end of the WebAssess stage of judging, you will have a set of VERY-USEFUL documents that are actually worth seeing, a set of USEFUL documents which aren't as good but at least answer the question, and a set of NOT-USEFUL documents which don't even answer the question.

Then, in the preference judging, you'll be making a pass over the VERY-USEFUL documents that contain a correct answer (we know what the answer actually is). The preference process helps identify the best-of-the-best, to put an ordering among the most useful documents.