Name: Medhanie Weldemariam

LiU-ID: medwe277

Email: medwe277@student.liu.se

1a. Explain recall with your own words.

It is one of the basic performance measures used in evaluating search strategies and Information retrieval: it can be expressed as the ratio of relevant documents correctly identified as relevant over all relevant documents including false negatives documents. Assume there is a database with a set of records and those records can be either relevant or irrelevant to what we are looking. Having this assumption in mind, we can define RECALL as the ratio of the number of relevant records retrieved to the total number of relevant records in the database. It is usually expressed as a percentage.

Thus, recall shows that out of the entire database collection, what proportion of all the relevant documents have been retrieved.

1b. Explain precision with your own words.

Precision is also one of the basic performance measures that we used in evaluating search strategies and Information retrieval. Let us consider the assumption that we used above in 1a. (RECALL). From this we can say that precision is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. This is also expressed as a percentage.

Thus, Precision shows from the retrieved documents, what proportion of them is relevant.

2a. Explain how recall can be decreased and increased using MeSH-terms in the created literature database.

In general, using full text and summary in information retrieval results in increasing recall. Since recall is inversely proportional to precision, it will lower precision.

According to our created literature database, recall can be increased by typing the truncation function/select all (*) to search for a records. The opposite way will decrease recall, in that using database boolean operators such as AND, NOT but not OR will greatly decrease recall.

2b. Explain how precision can be decreased and increased using MeSH-terms in the created literature database.

Both recall and precision vary inversely with each other, with greater precision (i.e., retrieving fewer irrelevant items) coming at the expense of finding fewer of the relevant items that exist in the database as a whole. Thus, MEDLINE citations show the highest precision, while full text has the highest recall.

Similarly when we come into our literature database, using boolean operators such as AND, NOT will highly increase precision but operators such as OR and (*) will decrease precision.

3a. Explain how other methods than MeSH-terms can be used to decrease and increase recall in the created literature database.

AND: Combining terms with AND decreases recall in that it lowers the number of search results, but increases the precision of the search. All the search terms must be present in each search result.

NOT: Combining terms with NOT decrease recall i.e. the number of results greatly decrease. But this will increase the precision of the search.

OR: Combining terms with OR increases recall (the number of results), but decreases the precision of the search. At least one of the search terms must be present in each search result.

3b. Explain how other methods than MeSH-terms can be used to decrease and increase precision in the created literature database.

AND: Combining terms with AND increases the precision of the search results. All the search terms must be present in each search result.

NOT: Similarly, Combining terms with NOT will increase precision of the result.

OR: Combining terms with OR increases the number of results. I.e recall, but decreases the precision of the search. At least one of the search terms must be present in each search result.

4. Does it look like there is a connection between recall and precision?

Yes, there is a connection in between, as I have tried to explain above, both are inversely related. As recall increases, precision will decrease and visa versa.