COMP3009J Information Retrieval
Worksheet 7: IR Evaluation - Programming

An Information Retrieval system has returned a set of results in response to a query. These results are contained in the following two files that you can download from Brightspace:
- **results.txt** contains the ranked list of results that an IR system has returned for a query.
- **qrels.txt** contains the relevance judgments for this query.

The files are in the format used in the Text REtrieval Conference (TREC), as follows:

The qrels file has 4 fields on each line, with a space between each:

`<query id> <iteration> <document id> <relevance>`

The "iteration" field is usually ignored and has the value 0. For these judged relevant documents, the level of relevance is included after each (0 means that it has been judged non-relevant). Any document not included in this file has not been judged.

The results file has 6 fields on each line, with a space between each:

`<query id> <iteration> <document id> <rank> <score> <run id>`

The important fields here are the query id, the document id and the rank (although the file is sorted by rank).

Write a program that reads the results and relevance judgments files, and calculates the following evaluation metrics. The correct answer is shown beside each.

- Precision (0.3333)
- Recall (0.75)
- Precision@10 (0.3)
- R-Precision (0.375)
- MAP (0.3132)
- bpref (0.4844)
- NDCG@10 (0.376)

Advanced
Your program should be able to calculate these metrics for files that have more than one query in them.