

Opinion-based QA over financial data #####
#####

- General Information:

The FiQA 2018 open challenge test collection is based on the use of unstructured text documents from different financial open data sources in English (microblogs, reports, news).

The FiQA task 2 data source contains information about questions and comments/posts extracted from Stackexchange, Stocktwits and Reddit.

- What is the content of each file?

The task 2 test collection consists of three files:

- FiQA question.tsv;
- FiQA question-doc.tsv;
- FiQA doc.tsv.

1- FiQA_train_question_final.tsv: This file contains a set of natural language questions on topics related to finance and investment. It consists of a .tsv file (tab-separated).

Columns:

- qid: question id number;
- question: question text;
- timestamp: the date and time value that represents when the question was posted.

2- FiQA_train_question_doc_final.tsv: This file contains information about the corresponding question-answer matchings. It also consists in a .tsv (tab-separated).

Columns:

- qid: question id number;
- docid: document id number.

3- FiQA_train_doc_final.tsv: This file contains a set of answers and comments that the systems need to find the matching one.

Columns:

- docid: document id number;
- doc: document text;
- timestamp: the date and time value that represents when the document was posted.