

# IR - Assignment 2 (Ranked retrieval)

## PART 1 (Ground Truth Preparation)

### Dataset Description:

1. query.txt contains total 82 queries, which has 2 columns: query id and query.
2. alldocs.rar contains documents file named with doc id. Each document has set of sentences.
3. output.txt contains 50 relevant documents (doc id) for each query

Link: [https://drive.google.com/file/d/1LOW6HJE\\_Y7lftHg58Zqccq70c\\_X78bh5/view?usp=sharing](https://drive.google.com/file/d/1LOW6HJE_Y7lftHg58Zqccq70c_X78bh5/view?usp=sharing)

## PREPARING GROUND TRUTH

For each query, rank the relevant 50 documents given by manually reading them and based on your judgement.

### Deliverables:

Query id	Ranked docID
Q1	<most relevant document ID>
Q1	<second most relevant document ID>

Name your file as A2\_part1<roll\_number>.txt

The students are allocated to the queries (ids are 1-82) according to the following:

<https://docs.google.com/spreadsheets/d/1nQrf2TRKvQo4b2EyEMnM1TzDI9EMnAY9M0yT1noCZRY/edit?ts=5b5bf744#gid=0>

## Deadline

30.08.2018