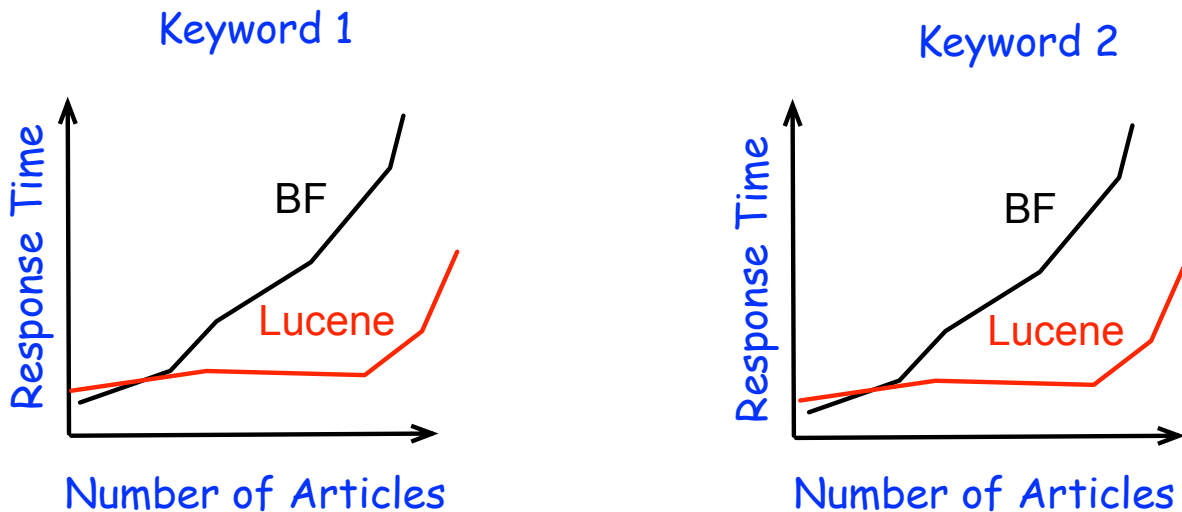**Homework 3**

Download the data from DBLP user: https://dblp.org/faq/1474679.html
Since the datafile is too big, you can download a small part (>5MB data) of it and not all of it.

1- Create a Lucene index for each of its data, key word: "Conversational Agent", "Chatbot", "Personal Assistant", "Smart Speaker"

2- Query the dataset for each of these keywords "Conversational Agent", "Chatbot", "Personal Assistant" and "Smart Speaker".

3- Plot the "response time" of the brute force search vs. Lucene ones for a keyword search, with different numbers of articles, including 1 article, 5 articles, 10 articles, 20 articles, 30 articles, …, 100 articles. Something like the following Figure, one for each keyword

Evaluation
Your homework grade will be calculated based on the following criteria:
Does the code compile? If not, you cannot get its mark.
Does it perform all tasks that have been asked in the assignment?
Is the code readable, did you put enough comments, respecting java naming conventions, etc.
Does it include a descriptive *read.me* file.
Are you able to explain your code and answer my questions about this? Cheating will not be tolerated at any cost.
- Any delay on assignment submission will result in grade deduction, 10% per day. After two days later the assignment will not get accepted at all.

- If there is an extreme circumstances such as getting sick or death of relatives, you can communicate with me about the late assignment. However, I would not accept anything without proper evidence.