

Search Engine Performance Module:

We created a class that writes the time taken for processes in text files. When the code is run, it will automatically generate those files.

First CrawlerPerformance.txt:

Time to crawl 1000 links is 121 seconds.

Time to crawl 100 links is 14 seconds.

Second: PageRankerPerformance.txt:

Time to page rank 1000 documents for 10 iterations is 461 seconds.

Time to page rank 100 documents for 10 iterations is 23 seconds.

The correct number of iterations should be equal to NumberOfDocuments - 1, but due to hardware limitations we chose to run 10 iterations instead.

Third: IndexerPerformance.txt:

Time to index and normalize 1000 documents is 830 seconds.

Time to index and normalize 100 documents is 47 seconds.

Fourth: QueryProcessor.txt

At 1000 links crawled:

Time to relevance rank 150 documents for a 2 words query is : 3 milliseconds.

Total latency to return search results is 245 milliseconds.

Total latency to return image results is 593 milliseconds.

Total latency to return trends 76 milliseconds.

At 100 links crawled:

Time to relevance rank 13 documents for a 2 words query is : 1 milliseconds.

Total latency to return search results is 75 milliseconds.

Total latency to return image results is 74 milliseconds.

Total latency to return trends 73 milliseconds.