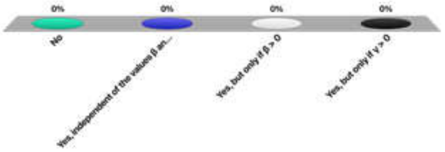
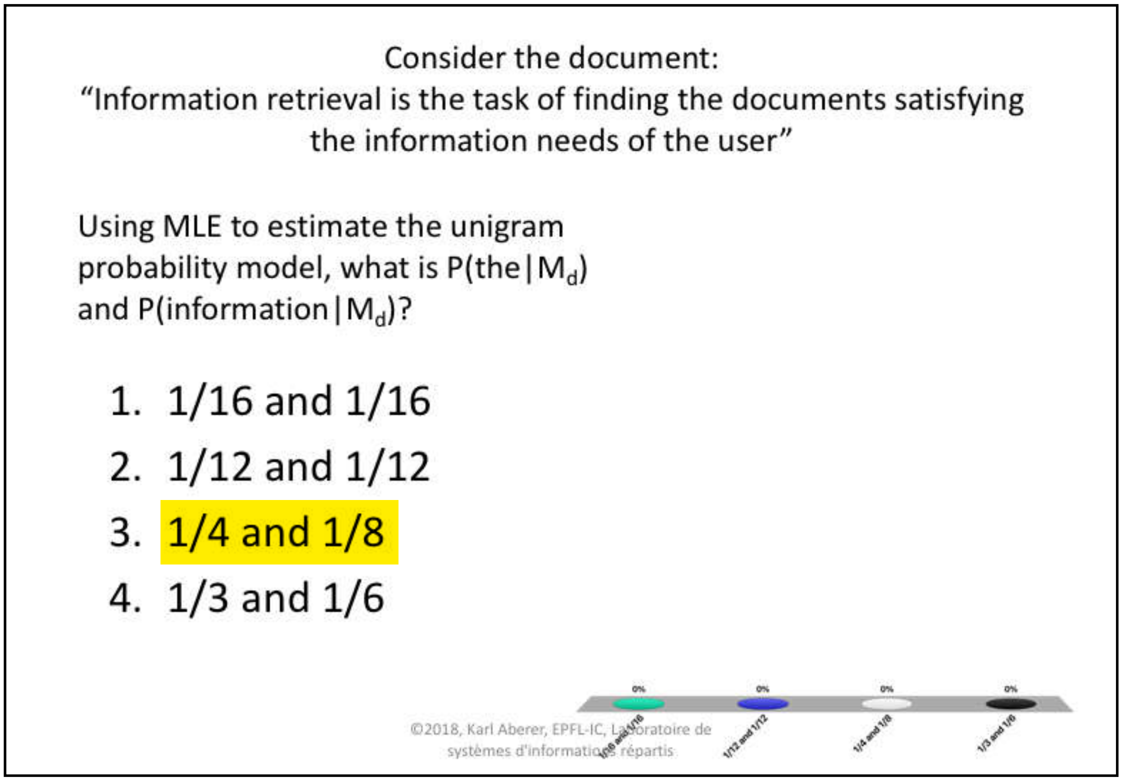
Can documents which do not contain any keywords of the original query receive a positive similarity coefficient after relevance feedback?

1. No
2. Yes, independent of the values β and γ
3. [✓ Yes, but only if β > 0]
4. Yes, but only if γ > 0





Consider the document:
“Information retrieval is the task of finding the documents satisfying the information needs of the user”

Using MLE to estimate the unigram probability model, what is P(the|M$_d$) and P(information|M$_d$)?

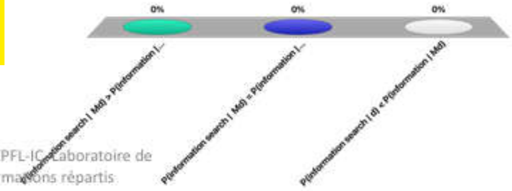
1. 1/16 and 1/16
2. 1/12 and 1/12
3. 1/4 and 1/8
4. 1/3 and 1/6

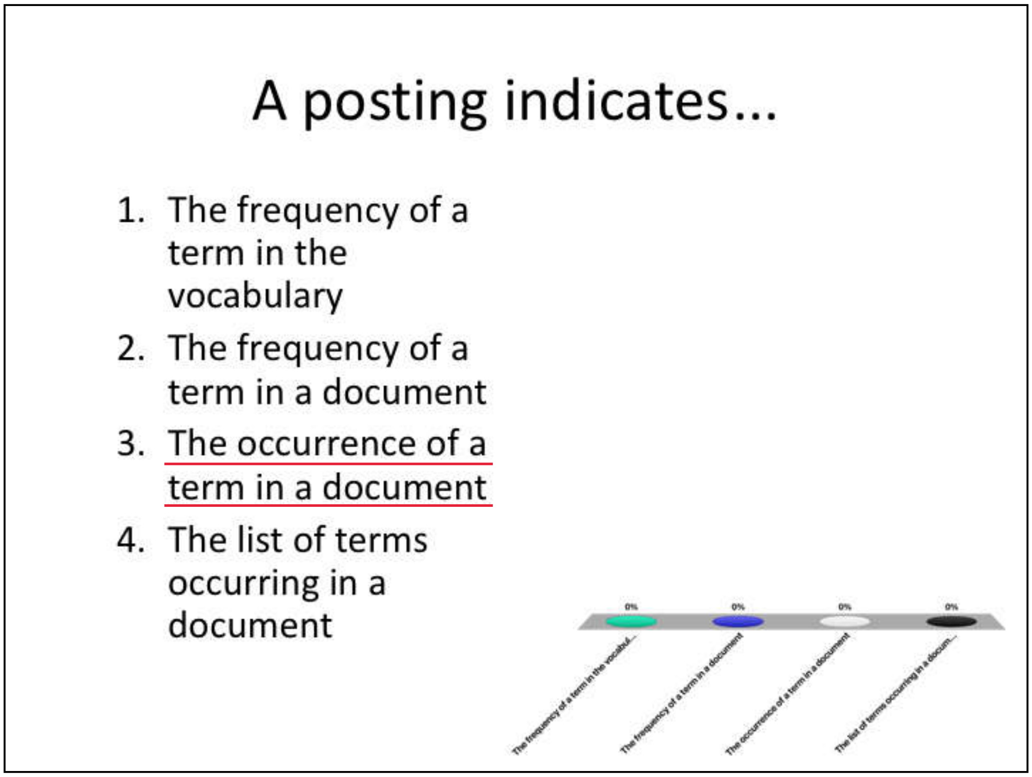
0%
0%
0%
0%
1/16 and 1/16
1/12 and 1/12
1/4 and 1/8
1/3 and 1/6

Consider the following document

d = “information retrieval and search”

1. P(information search | M_d) > P(information | M_d)
2. P(information search | M_d) = P(information | M_d)
3. P(information search | d) < P(information | M_d)

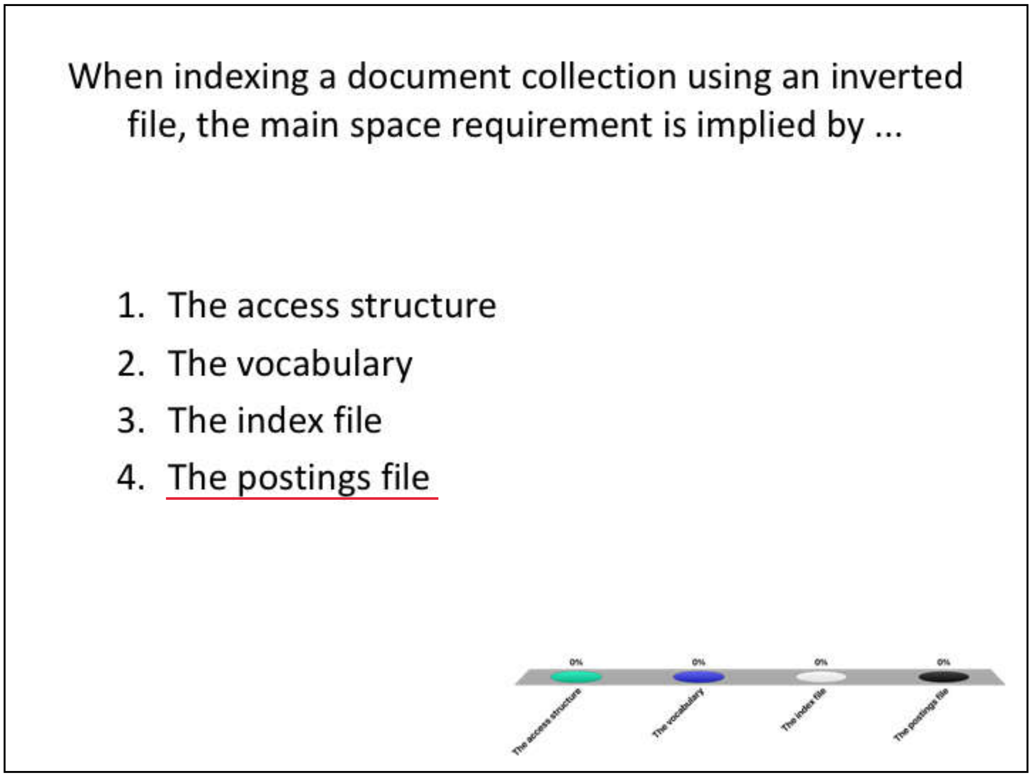




A posting indicates...

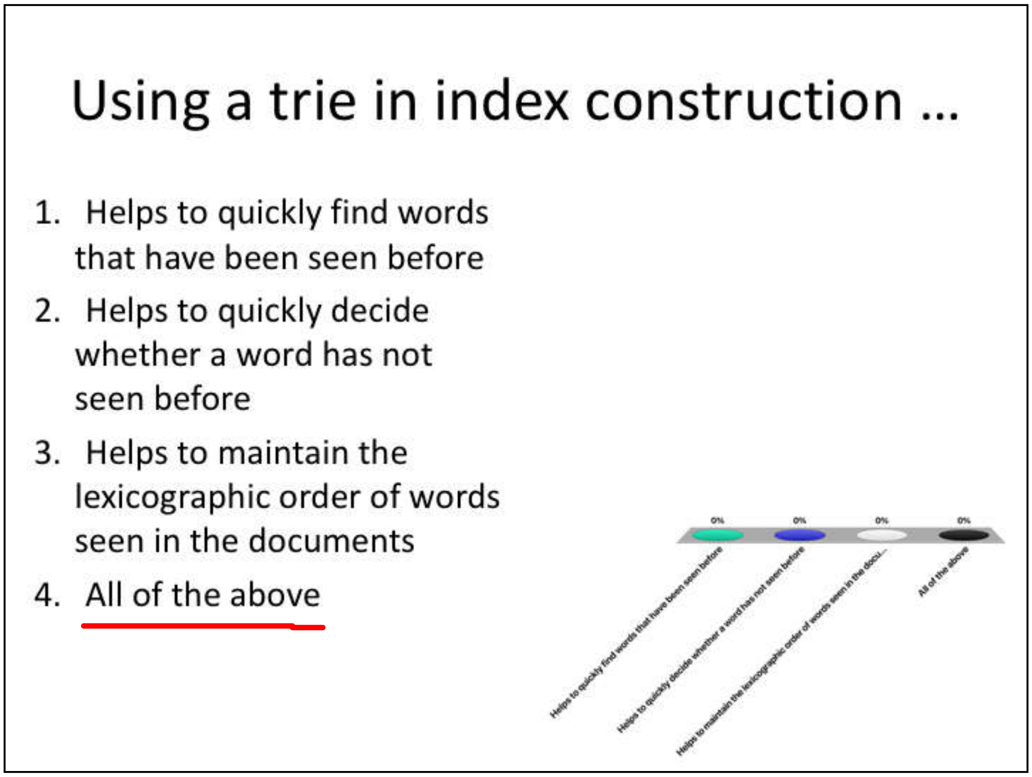
1. The frequency of a term in the vocabulary
2. The frequency of a term in a document
3. The occurrence of a term in a document
4. The list of terms occurring in a document

0%
0%
0%
0%
The frequency of a term in the vocabul...
The frequency of a term in a document
The occurrence of a term in a document
The list of terms occurring in a docum...



When indexing a document collection using an inverted file, the main space requirement is implied by ...

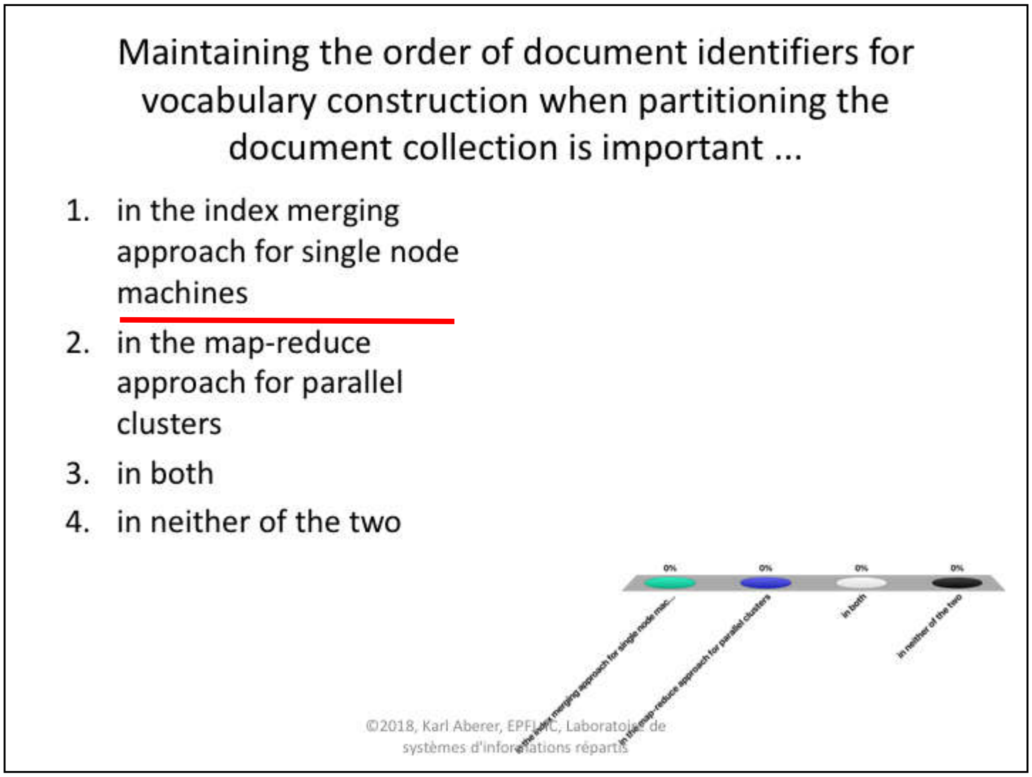
1. The access structure
2. The vocabulary
3. The index file
4. The postings file



Using a trie in index construction …

1. Helps to quickly find words that have been seen before
2. Helps to quickly decide whether a word has not seen before
3. Helps to maintain the lexicographic order of words seen in the documents
4. All of the above

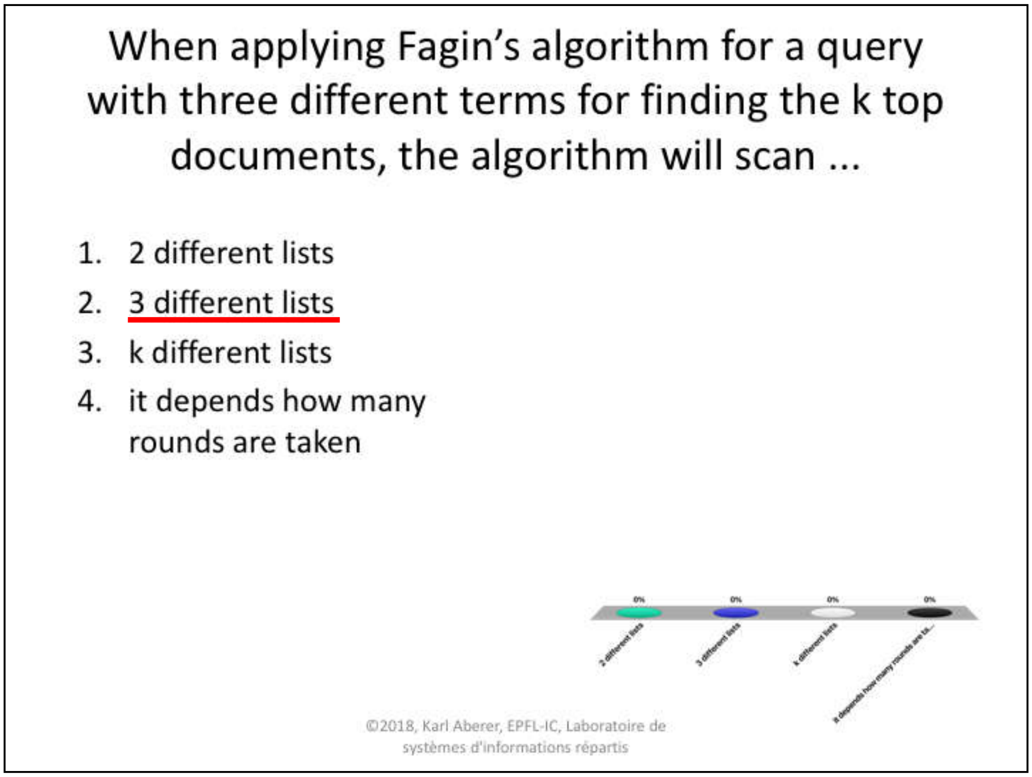
0%
0%
0%
0%
Helps to quickly find words that have been seen before
Helps to quickly decide whether a word has not seen before
Helps to maintain the lexicographic order of words seen in the docu...
All of the above



Maintaining the order of document identifiers for vocabulary construction when partitioning the document collection is important ...

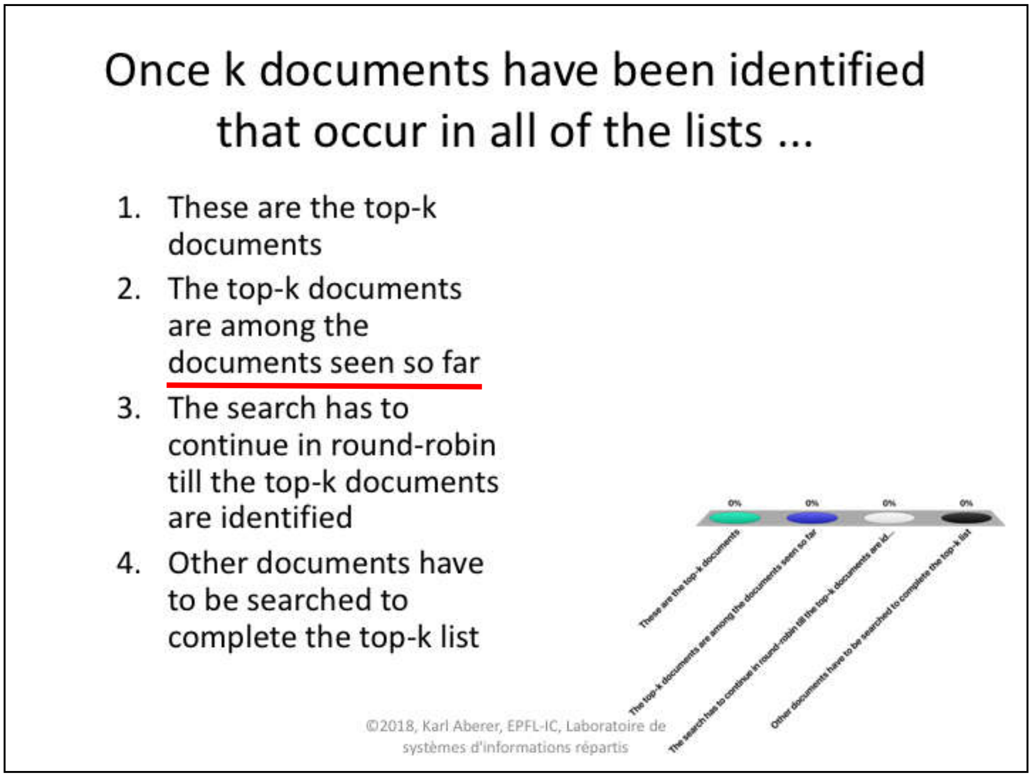
1. in the index merging approach for single node machines
2. in the map-reduce approach for parallel clusters
3. in both
4. in neither of the two

0%
0%
0%
0%
in the index merging approach for single node mac...
in the map-reduce approach for parallel clusters
in both
in neither of the two



When applying Fagin's algorithm for a query with three different terms for finding the k top documents, the algorithm will scan ...

1. 2 different lists
2. [✓ 3 different lists]
3. k different lists
4. it depends how many rounds are taken



Once k documents have been identified that occur in all of the lists ...

1. These are the top-k documents
2. The top-k documents are among the documents seen so far
3. The search has to continue in round-robin till the top-k documents are identified
4. Other documents have to be searched to complete the top-k list

0%
0%
0%
0%
These are the top-k documents
The top-k documents are among the documents seen so far
The search has to continue in round-robin till the top-k documents are id...
Other documents have to be searched to complete the top-k list