A retrieval model attempts to capture …

1. the interface by which a user is accessing information
2. the importance a user gives to a piece of information for a query
3. the formal correctness of a query formulation by user
4. the structure by which a document is organised

Full-text retrieval refers to the fact that …

1. the document text is grammatically fully analyzed for indexing
2. queries can be formulated as texts
3. all words of a text are considered as potential index terms
4. grammatical variations of a word are considered as the same index terms

The entries of a term-document matrix indicate …

1. how many relevant terms a document contains
2. how frequent a term is in a given document
3. how relevant a term is for a given document
4. which terms occur in a document collection

Let the query be represented by {(1, 0, -1), (0, -1, 1)} and the document by (1, 0, 1). The document …

1. matches the query because it matches the first query vector
2. matches the query because it matches the second query vector
3. does not match the query because it does not match the first query vector
4. does not match the query because it does not match the second query vector

The term frequency of a term is normalized …

1. by the maximal frequency of all terms in the document
2. by the maximal frequency of the term in the document collection
3. by the maximal frequency of any term in the vocabulary
4. by the maximal term frequency of any document in the collection

The inverse document frequency of a term can increase …

1. by adding the term to a document that contains the term
2. by removing a document from the document collection that does not contain the term
3. by adding a document to the document collection that contains the term
4. by adding a document to the document collection that does not contain the term

If the top 100 documents contain 50 relevant documents …

1. the precision of the system at 50 is 0.25
2. the precision of the system at 100 is 0.5
3. the recall of the system is 0.5
4. All of the above

If retrieval system A has a higher precision at k than system B …

1. the top k documents of A will have higher similarity values than the top k documents of B
2. the top k documents of A will contain more relevant documents than the top k documents of B
3. A will recall more documents above a given similarity threshold than B
4. the top k relevant documents in A will have higher similarity values than in B

Let the first four documents retrieved be R N N R. Then the MAP is

1. 1/2
2. 3/4
3. 2/3
4. 5/6