```
import re
from collections import defaultdict


document = {1:'This is the fisrt document. It contains some text.',
            2: 'The second document is longer. It also contains some text.',
            3:'This is the third document. It is different from the first two'}


def preprocess_doc(doc):
  tokens = re.findall(r'\w+', doc.lower())
  stopwords = set(['the','is','it','and','some'])
  tokens = [token for token in tokens if token not in stopwords]
  return tokens


inverted_index = defaultdict(list)

for doc_id, doc_text in document.items():
  tokens = preprocess_doc(doc_text)
  for token in tokens:
    inverted_index[token].append(doc_id)


def retrieve_doc(query):
  query_tokens = preprocess_doc(query)
  result = set()
  for token in query_tokens:
    if token in inverted_index:
      result.update(inverted_index[token])
  return list(result)


query = 'document contains text'
matching_doc = retrieve_doc(query)


if matching_doc:
  print('Matching document for query:',query)
  for doc_id in matching_doc:
    print('Document',doc_id,':',document[doc_id])
else:
    print('No matching documents found')
```

```
Matching document for query: document contains text
Document 1 : This is the fisrt document. It contains some text.
Document 2 : The second document is longer. It also contains some text.
Document 3 : This is the third document. It is different from the first two
```

Start coding or generate with AI.