

```

document1 = "The quick brown fox jumped over the lazy dog."
document2 = "The lazy dog slept in the sun."

# Step-1: Tokenize the documents

tokens1 = document1.lower().split()
tokens2 = document2.lower().split()

print("tokens 1: ", tokens1)
print("tokens 2: ", tokens2)

tokens 1:  ['the', 'quick', 'brown', 'fox', 'jumped', 'over', 'the',
'lazy', 'dog.']
tokens 2:  ['the', 'lazy', 'dog', 'slept', 'in', 'the', 'sun.']

# Combine the tokens into the list of unique terms

terms = list(set(tokens1 + tokens2))
print(terms)

['the', 'in', 'slept', 'jumped', 'brown', 'quick', 'over', 'dog.',
'dog', 'sun.', 'lazy', 'fox']

# Step-2: Build the inverted index
inverted_index = {}

# For each term find the document that contains it

for term in terms:
    document = []
    if term in tokens1:
        document.append('Document 1')
    if term in tokens2:
        document.append('Document 2')
    inverted_index[term] = document
    print(document)

['Document 1', 'Document 2']
['Document 2']
['Document 2']
['Document 1']
['Document 1']
['Document 1']
['Document 1']
['Document 1']
['Document 1']
['Document 2']
['Document 2']
['Document 1', 'Document 2']
['Document 1']

```

```
for term, document in inverted_index.items():
    print(term,"->", ",".join(document))

the -> Document 1,Document 2
in -> Document 2
slept -> Document 2
jumped -> Document 1
brown -> Document 1
quick -> Document 1
over -> Document 1
dog. -> Document 1
dog -> Document 2
sun. -> Document 2
lazy -> Document 1,Document 2
fox -> Document 1
```