[Day-32 2211cs020196] Write a Python program to load a text file, perform tokenization, calculate the term frequency (TF) of each token, and display the top 5 most frequent tokens.

In [1]:
```python
import nltk
from collections import Counter
import re
nltk.download('punkt')
def tokenize_text(text):
    tokens = nltk.word_tokenize(text)
    return tokens
def calculate_tf(tokens):
    tf = Counter(tokens)
    return tf
text = """
The quick brown fox jumps over the lazy dog. The quick brown fox is very quick and very brown.
The dog, however, is just lazy. Foxes are generally quick and dogs are sometimes lazy.
"""
text = re.sub(r'\W+', ' ', text).lower()
tokens = tokenize_text(text)
tf = calculate_tf(tokens)
top_5_tokens = tf.most_common(5)
print("Top 5 Most Frequent Tokens:")
for token, freq in top_5_tokens:
    print(f"{token}: {freq}")

```

```
[nltk_data] Downloading package punkt to C:\Users\Sai Krishna
[nltk_data]     Hari\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!

Top 5 Most Frequent Tokens:
the: 4
quick: 4
brown: 3
lazy: 3
fox: 2
```