

TiktokenText Splitter

1. How the text is split: by `tiktoken` tokens
2. How the chunk size is measured: by `tiktoken` tokens

```
# This is a long document we can split up.  
with open('../..../state_of_the_union.txt') as f:  
    state_of_the_union = f.read()
```

```
from langchain.text_splitter import TokenTextSplitter
```

```
text_splitter = TokenTextSplitter(chunk_size=10, chunk_overlap=0)
```

```
texts = text_splitter.split_text(state_of_the_union)  
print(texts[0])
```

Madam Speaker, Madam Vice President, our