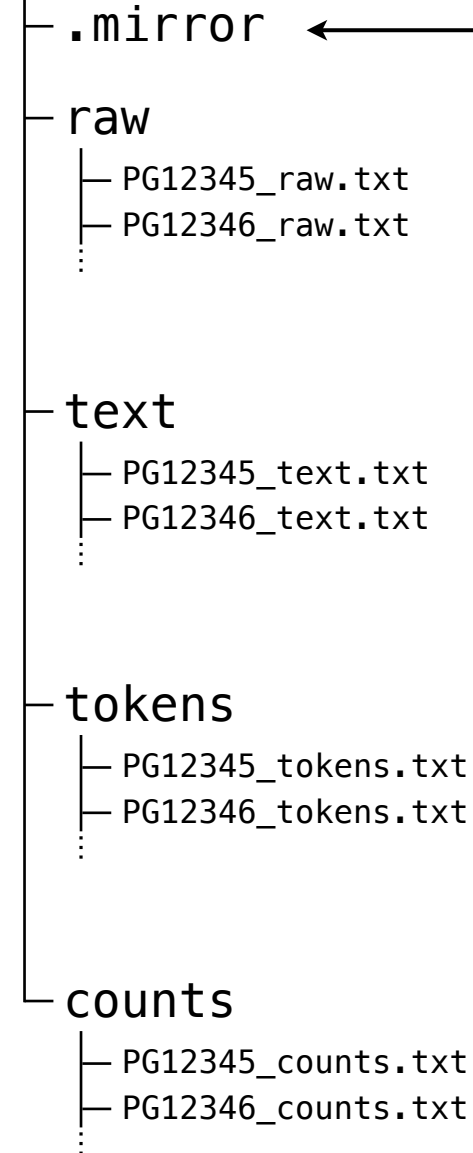


data



\$ **rsync** aleph.gutenberg.org::gutenberg

Header

A long time ago
in a galaxy
far, far away...

Tail

A long time ago
in a galaxy
far, far away...

a
long
time
ago
in
...

the 3245
of 1554
:
galaxy 123
long 87
far 34

> **nlTK.tokenize.sent_tokenizer**
> **nlTK.[...].TreebankWordTokenizer**
> **string.isalpha()**
> **string.lower()**