lexical analysis

source .tex file

pos = 23

pos = 24

```latex
\documentclass{article}
\title{The Declaration of Independence}
\author{Thomas Jefferson}
\date{June 28, 1776}
\begin{document}
\maketitle
   When in the Course of human events, it
becomes necessary for one people to
disolve the poiltical bands which have
connected them with another, the separate
and equal station to which the Laws of
Nature and of Nature's God entitle them,
a decent respect to the opinions of
mankind requires that they should declare
the causes which impel them to the
separation.

   We hold these truths to be self-
evident, that all men are created equal,
that they are endowed by their Creator
with certain unalienable Rights, that
among these are Life, Liberty and the
pursuit of Happiness.
\end{document}
```

tokenizing

Token	Meta data
⬜	START_OF_FILE @ 0 (0)
\documentclass	COMMAND @ 0 (14)
{	SYMBOL @ 14 (1)
article	WORD @ 15 (7)
}	SYMBOL @ 22 (1)
\title	COMMAND @ 24 (6)
	...
\date	COMMAND @ 91 (5)
{	SYMBOL @ 96 (1)
June	WORD @ 97 (4)
28	WORD @ 102 (2)
,	PUNCTUATION @ 104 (1)
1776	WORD @ 106 (4)
	...
\separation	WORD @ 492 (10)
.	PUNCTUATION @ 502 (1)
⬜	PARAGRAPH @ 503 (2)
We	WORD @ 507 (2)
hold	WORD @ 510 (4)
	...
document	WORD @ 722 (8)
}	SYMBOL @ 730 (1)
⬜	END_OF_FILE @ 732 (0)

token

meta data such as
type, position and
length of lexeme

specialized (empty) tokens for
paragraphs, start and end of file