This sentence is expected to be extracted as a single line because the user hasn’t hit any line return on the keyboard, keeping the line return in added by the editor for visualization will make NER mode complicated

In contrast this once should appear on a new line

And same for this one which should ideally be separated from the previous one by a blank line

Note: by copy pasting the content into a text editor, it will appear as described upper, where the output of Tika contains actual line break at the end of PDF lines