

## PDF parsing

1. Save a wordfile as PDF, open it in Python, extract all text. Describe the differences, if any. Try the same with a two column PDF file from the web. This exercise gets more interesting if you use `_difficult_` PDF. Why not try <http://wch.github.io/latexsheet/latexsheet.pdf>

- Is the word order still as it should be?
- What about the strange characters?