

Document title

This is a *simple **document*** in PDF format.

This is a paragraph that spans three lines. These should be treated as three textboxes in the document object model, but as a single box by the textboxes() method.

This is a paragraph that is set entirely in bold. Pdftohtml will create a element for the text in this paragraph.

This paragraph leads with bold. But then it switches to regular. And it spans more than one line, which requires a tricky merge of TextElement objects.