

Hand-coded PDF sample files

This directory hosts a few hand-coded PDF sample files.

They are designed to make it easy playing with them in a text editor. Most of them contain commented lines explaining some technical background about the PDF features they use. Inline comments also include a few suggestions about possible experiments to conduct with them.

Before you play with these PDFs and modify them, create backups!

- **102_A-vectorized.pdf :**
Looks like glyph for character ‘A’, but is a vector shape. Text cannot be copied or extracted. Play with drawing color, line segments, metadata entries and `cm` operator.
- **104_fonts-not-embedded.pdf :**
Uses “real” text with font, but font is not embedded. Text can be copied or extracted. Play with text color, font size, font used, page size, page rotation...
- **105_transformation-matrix.pdf :**
Play with the `cm` (concatenation matrix) operator, applied to text.
- **106_hello-troopers.pdf :**
Play with RGB colors and with the `cm` operator, applied to text.
- **107_perspectives-by-banksy.pdf :**
PDF contains one image. Play with parameters for the dimensions of the image. Also play with `/CropBox` and `/Rotate` keys of the page dictionary.
- **108_text-rendering-modes.pdf :**
PDF contains text. Change PDF page size. Change text color to “white”, using RGB color values. Extract or copy text from PDF page. Text may be rendered in different modes. Play with the `Tr` operator.
- **111_current-transformation-matrix-ctm.pdf :**
PDF contains one 2x2 pixels image. Image is re-used multiple times on the page, each time at different location, with different scaling/skew/rotation by using different `cm` values. Observe the effects of enabling/disable different `cm` values (by commenting in/out respective PDF source lines). Open file in different PDF viewers. Do they all render the 2x2 image identically?
- **112_play-with-tounicodetable.pdf :**
2 pages, (almost) identical content. Second page’s content stream is slightly obfuscated. Experiment with the `/Widths` array for a font used by the file. Also play with the value for the `/FirstChar` key. Study the influence of the `/ToUnicode` table upon the text extraction success. 3 different `/ToUnicode` tables are included. You can enable one of them (or none) and observe the effects upon text extraction or text copy’n’paste capabilities.
- **113_stegano-with-tounicodetable.pdf :**
A variation of the `112_play-with-tounicodetable.pdf`. See how a manipulation of the `/ToUnicode` table can in effect apply a `rot13`-like effect when extracting text.
- **114_incrementally-updated.pdf :**
Delete all lines after the first `%%EOF` and observe the effects on the file. Test also with `pdftotext` or with copy’n’pasting of text.
- **115_little-riddle.pdf :**
A littel riddle: Something is hidden in this PDF – what is it? How did the hiding happen??