

В сторону анализа содержимого входного и выходных файлов, вестимо. Примеры файлов покажите.

Вообще, если есть текстовое содержимое, в какой-то степени так и должно быть. Поскольку при нарезке на страницы в каждый отдельный файл будут встроены наборы шрифтов, которые в исходном файле присутствуют в одном экземпляре. Но это может усугубляться иными причинами. Например, если шрифты попутно преобразовываются в `type3`, размер файла может вырасти в сотню раз. С картинками и прочими объектами тоже есть нюансы.

Всем привет!

Код: [Выделить]

```
$ du -b file.pdf
15820146      file.pdf
```

Код: [Выделить]

```
$ pdftk A=file.pdf cat A1 output 1.pdf
```

Код: [Выделить]

```
$ du -b 1.pdf
15654155      1.pdf
```

Код: [Выделить]

```
$ pdftk A=file.pdf cat A3-7 output 5.pdf
```

Код: [Выделить]

```
$ du -b 5.pdf
15668428      5.pdf
```

При этом в файлах 1.pdf и 5.pdf реально по 1 и 5 страниц соответственно.

pdfseparate также раздувает каждую страницу, она всегда практически равна объёму исходного файла.

Куда копать?

p9