*Язык: Java*

*Середа: SublimeText*

*Использование Apache POI для разбора документов Word*

*В данной работе описывается как скопировать часть текстового docx файла в другой docx файл, с сохранением используемых стилей. Задача кажется тривиальной, но на практике возникли некоторые трудности.  
     Итак, предположим, есть большой docx файл, состоящий из логически разделенных частей, разделенных, скажем, разрывом страницы (Ctrl+Enter). Требуется создать на каждый такой кусок отдельный файл, с сохранением форматирования текста.  
    Воспользуемся библиотекой Apache POI  для разбора файла. Для этого создадим проект и добавим в зависимости POI 3.8. Версия 3.7 не подходит по причине того, что в ней нет возможности перенести стили из одного документа в другой.*

*См.файл Apachepoidocfilse.java*

*Удобнее генерировать на doc/docx, а rtf/xml/html*

*См.фaйл generalDoxDocx.java*

*Чтение из doc Word файла с использованием Apache POI.  
В примере я считываю два типа информ.*

*Все работы осуществляются через чтение****Paragraph****, но вот проблема - в ячейках таблиц как и в обычном тексте содержатся****Paragraph****-ы. Потому сортирую с помощью****if (!par.isInTable())****на те что содержатся в таблицах и на те что не содержатся. i++ использовать приходится чтобы отсечь****Paragraph****-ы которые содержатся в таблицах. По причине того, что каждая строка таблицы завершается пустым****Paragraph****-ом, то их тоже надо отсекать, для этого пришлось добавить****i++****и в переборе цикла****TableRow****.ации текст и таблицы*

*См.файл ApachePoidoc.java*