**Назначение.**

Утилита для «сжатия» файлов в формате PDF, предназначена для изменения разрешения встроенных в документ изображений, документы, при этом, должны находится в системе RSDOC.

Утилита представляет собой приложение для командной строки написанное на языке Java, параметры работы задаются при помощи файла настроек.

**Файлы и директории.**

Программа не требует установки и каких-либо библиотек кроме развёрнутой виртуальной машины Java версии 8, после разархивирования необходимо отредактировать файл настроек pdfpack.cfg, образец находится в pdfpack.cfg.template

**Работа программы**

После запуска считывается файл настроек, по указанному в нём адресу с использованием заданных имени пользователя и пароля осуществляется попытка подключения к базе данных системы RSDOC,в случае удачного подключения последовательно считываются документы с заданными ID (список ID задаётся в файле), для перечисленных документов из базы в резервную папку скачиваются файлы документов. Если в резервной папке уже имеется файл с идентификатором документа, работа программы может быть остановлена будет она остановлена или нет зависит от настроек. Каждый файл документа обрабатывается следующим образом :

1. Из документа во временную папку извлекаются все встроенные файлы образов
2. Параллельно запускаются несколько потоков которые пытаются изменить разрешение каждого извлечённого файла в соответствии с настройками и ограничениями
3. Если файл уже был упакован, он содержит метаданные со сводкой и датой сжатия – такой файл будет пропущен
4. Файлы с изменённым разрешением сохраняются на диск
5. В резервную папку сохраняется новый файл в формате PDF в котором некоторые изображения заменяются на новые версии с изменённым разрешением
6. В новый файл добавляются метаданные содержащие сводку с информацией о сжатии и коротким резюме для каждого файла
7. Сравнивается размер вновь созданного файла с изначальным, если размер уменьшился, новый файл сохраняется в базу RSDOC (сохранение в базу может быть отключено в файле настроек)
8. Временные файлы удаляются

Содержимое резервной директории остаётся неизменным.

Программа сохраняет журнал своей работы в котором регистрируется сжатие изображений и выполненные действия в базе данных.

**Параметры конфигурации.**

Параметры конфигурации считываются из XML файла pdfpack.cfg, пример конфигурации находится в файле pdfpack.cfg.template параметры конфигурации :

|  |  |
| --- | --- |
| **Параметр (ТЕГ)** | **Назначение** |
| ignore | Игнорировать изображение если сжатие даёт результат в %\*10 меньше чем значение параметра. Т.е. если значение параметра равно 10 изображение не будет заменятся если результат сжатия меньше или равен 0.1% |
| dpi | DPI изображения до которого надо уменьшать от исходного |
| algo | Алгоритм сжатия (изменения размера картинки |
| minfile | Минимальный размер файла изображений при котором он исключается из обработки |
| minpic | Минимальный размер картинки в пикселях по горизонтали или вертикали при котором она исключается из обработки |
| tmp | Каталог для временных файлов |
| db | JDBC URL для подключения к базе данных |
| user | Пользователь для подключения к базе данных |
| pass | Пароль пользователя базы данных |
| docfile | Текстовый файл, в котором находятся ID документов из системы RSDOC |
| docsrc | Источник из которого будут поступать ID документов из системы RSDOC возможные значения :  «db» программа выполнится запрос (текст запроса находится в files.txt) в соот. базе данных и результаты этого запроса будут использоваться как ID файлов  «file» ID файлов в формате PDF будут браться из файла заданного параметром docfile |
| stoponerror | Логический параметр (возможные значения true или false) если установлен в true то при первой ошибке (сжатия или работы с БД) работа процесс обработки документов будет приостановлен |
| changedb | Логический параметр (возможные значения true или false) если установлен в false то в БД не будут вносится изменения |
| backup | Каталог для резервирования файлов |

Пример конфигурации для тестовой БД находится в файле pdfpack.cfg.template.

**Сценарий работы.**

1. Проверяем установлена ли Java версии 1.8 и есть ли доступ к серверу
2. Скачиваем скомпилированную версию программы, разворачиваем.
3. Делаем бэкап базы данных (рекомендуется)
4. Создаём конфигурационный файл pdfpack.cfg с нужным подключением к БД и другими параметрами (пример конфига есть в файле pdfpack.cfg.template)
5. Перед запуском программы очень желательно сделать резервную копию БД
6. Теперь необходимо получить список обрабатываемых документов сделать это можно при помощи запроса files.sql, имеет смысл обрабатывать файлы порциями и наблюдать за полученным результатам по файлам журнала. Если установить параметр «docsrc» будет загружен и выполнен запрос из файла files.txt и все его результаты будут использоваться как ID файла в системе RSDOC