Оглавление

[1 Описание приложения 2](#_Toc440818548)

[2 Установка и конфигурация приложения 3](#_Toc440818549)

[2.1 Предварительные действия 3](#_Toc440818550)

[2.2 Содержимое поставки 3](#_Toc440818551)

[2.3 Установка приложения 3](#_Toc440818552)

[2.4 Конфигурация приложения 4](#_Toc440818553)

[2.4.1 Главный конфигурационный файл приложения 4](#_Toc440818554)

[2.4.2 Конфигурация формата отчетов 5](#_Toc440818555)

[2.4.3 Конфигурация правил сортировки 6](#_Toc440818556)

[2.4.4 Конфигурация типов электронной почты 6](#_Toc440818557)

[2.4.5 Шаблоны отчетов 7](#_Toc440818558)

[3 Работа с приложением 9](#_Toc440818559)

# Описание приложения

Данное приложение предназначено для автоматизации составления отёчности по присылаемым рефератам. Приложение реализует следующий бизнес-процесс.

1. Скачивание doc- и docx-файлов с почты на постоянную память компьютера.
2. Сортировку файлов по директориям, в зависимости от шаблона имени файла.
3. Составление отчетности по файлам для каждой директории сортировки с использованием содержимого файлов.

Исходный код приложения доступен по адресу: <https://github.com/androsovgr/document-reporter>

# Установка и конфигурация приложения

В данном разделе описана установка и настройка приложения.

## Предварительные действия

1. Необходимо скачать и установить java версии не ниже, чем 1.8. Последняя версия Java доступна по ссылке: <https://www.java.com/ru/download/>
2. Путь до Java версии 1.8 необходимо добавить в переменную среды PATH, если это не было сделано автоматически. Проверить можно, введя команду: «java - version» в командную строку – должна вернуться версия Java. Процесс изменения переменной среды PATH описан: <https://www.java.com/ru/download/help/path.xml>

## Содержимое поставки

В архиве с поставкой приложения содержатся следующие папки и файлы.

|  |  |
| --- | --- |
| Имя папки | Описание |
| configuration | Содержит конфигурационные файлы приложения |
| libs | Содержит библиотеки приложения |
| docs | Папка с документацией к приложению |
| document-reporter-ui-<ВЕРСИЯ>.jar | Библиотека – пользовательский интерфейс приложения. |
| Ui.bat | Файл запуска приложения. |

## Установка приложения

Содержимое архива нужно распаковать в любую директорию, к которой у пользователя приложения будут права на запись, чтение, исполнение bat-файлов. Для работы приложения нужно также определить «рабочую директорию» - директорию, в которую будут скачиваться рефераты с почты и сохраняться отчеты. Структура рабочей директории:

|  |  |
| --- | --- |
| Папка | Назначение |
| all | Сюда первоначально попадают все вложения с почты. В ней остаются нераспределенные вложения (не соответствующие ни одному правилу сортировки), пока пользователь их не удалит. |
| sorted | В этой папке располагаются другие папки, в которые попадают рефераты в соответствие правилам сортировки |
| report | В этой папке составляются отчеты |

## Конфигурация приложения

Приложение содержит следующие конфигурационные файлы

|  |  |
| --- | --- |
| Назначение | Место |
| Главный конфигурационный файл приложения | configuration/main.properties |
| Форматы составляемых отчетов и классы, которые для этого используются. | configuration/report-writers.xml |
| Правила сортировки скачиваемых файлов по папкам | configuration/sortings.xml |
| Шаблоны, в соответствие которым извлекаются данные из документов и строятся отчеты | Xml-файлы в директории configuration/doc-template |
| Описание правил подключения к различным почтовым хранилищам (gmail, mail.ru, yandex, др.) | Файлы с расширением properties в директории configuration/mail |

### Главный конфигурационный файл приложения

Файл «main.properties» должен быть валидным property-файлом: его содержимое должно быть заэкранировано.

|  |  |  |
| --- | --- | --- |
| Атрибут | Назначение | Пример значения |
| email.login | Адрес электронной почты | mephi\.essay@gmail.com |
| email.password | Пароль электронной почты | 123123 |
| email.type | Тип электронной почты (gmail, др.) Соотвествует имени property-файла в папке configuration/mail | google |
| work.folder | Папка, в которую выгружаются рефераты и в которой создаются отчеты | E:\\TMP\\report-work |
| work.lastload | Timestamp (дата в милисекундах от 1 января 1970 года). Время получения письма с последним выгруженным рефератом. С почты загружаются рефераты только присланные после этой даты. После получения очередного реферата дата обновляется. Для получения значения, соответствующего нужной дате, можно использовать онлайн-конвертеры (например, <http://www.epochconverter.com/>). | 1452538153000 |

### Конфигурация формата отчетов

XML-файл. Пример файла:

<?xml version=*"1.0"* encoding=*"UTF-8"*?>

<report-writers>

<report-writer>

<id>docx</id>

<extention>docx</extention>

<writer-class>ru.mephi.dr.writer.DocxTableWriter</writer-class>

</report-writer>

<report-writer>

<id>xlsx</id>

<extention>xlsx</extention>

<writer-class>ru.mephi.dr.writer.XlsxTableWriter</writer-class>

</report-writer>

<report-writer>

<id>xls</id>

<extention>xls</extention>

<writer-class>ru.mephi.dr.writer.XlsTableWriter</writer-class>

</report-writer>

</report-writers>

Описание содержимого файла.

|  |  |
| --- | --- |
| Путь до тэга | Описание |
| report-writers.report-writer.id | Идентификатор типа отчета. Отображается в приложении |
| report-writers.report-writer.extention | Расширение файла составляемого отчета |
| report-writers.report-writer.writer-class | Java-класс, отвечающий за запись информации в отчет. Информацию по всем классам можно найти в папке docs\apidocs\ru\mephi\dr\writer\class-use |

### Конфигурация правил сортировки

XML-файл. Пример файла:

<?xml version=*"1.0"* encoding=*"UTF-8"*?>

<sortings>

<sorting>

<id>history</id>

<name-pattern>history.\*\.docx?</name-pattern>

<folder>history</folder>

</sorting>

<sorting>

<id>phylosophy</id>

<name-pattern>phylosophy.\*\.docx?</name-pattern>

<folder>phylosophy</folder>

</sorting>

</sortings>

Описание содержимого файла.

|  |  |
| --- | --- |
| Путь до тэга | Описание |
| sortings.sorting.id | Уникальный идентификатор правила сортировки. Соответствует имени шаблона документа |
| sortings.sorting.name-pattern | Регулярное выражение, соответствующее шаблону имен файлов, попадающих под условия этого правила сортировки |
| sortings.sorting.folder | Имя папки внутри директории sorted рабочей директории, в которую будут перемещены документы, соответствующие правилу сортировки |

### Конфигурация типов электронной почты

Текущая версия приложения работает со следующими хранилищами электронной почты:

* gmail.

Для поддержки дополнительных хранилищ нужно добавить конфигурационные файлы в эту папку. Содержимое и даже свойства конфигурационного файла зависят от типа хранилища.

### Шаблоны отчетов

XML-файл. Пример файла:

<template xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

xsi:schemaLocation="template.xsd">

<attribute key="subject" label="Предмет"

retrieverClass="ru.mephi.dr.parser.retriever.CellValueRetriever"

silence="true">

<parameters>

<parameter>

<key>expression</key>

<value>Реферат по (.\*) аспиранта кафедры.\*</value>

</parameter>

<parameter>

<key>table-position</key>

<value>first</value>

</parameter>

<parameter>

<key>row-number</key>

<value>3</value>

</parameter>

<parameter>

<key>column-number</key>

<value>1</value>

</parameter>

<parameter>

<key>line</key>

<value>2</value>

</parameter>

</parameters>

<validators>

<validator validatorClass="ru.mephi.dr.parser.validator.ValueLikeValdator"

validatorMessage="Тема реферата неправильная">

<parameters>

<parameter>

<key>regexp</key>

<value>истори.? наук.?</value>

</parameter>

</parameters>

</validator>

</validators>

</attribute>  
<template>

|  |  |
| --- | --- |
| Путь до тэга | Описание |
| template.attribute | Информация, извлекаемая из реферата (ФИО, номер кафедры, др.). Один аттрибут соответствует одному столбцу отчета |
| template.attribute.key | Уникальный идентификатор аттрибута |
| template.attribute.label | Заголовок колонки с атрибутом в отчете |
| template.attribute.retrieverClass | Java-класс, отвечающий за получение аттрибута из документа определенным образом. Документация по всем таким классам – в папке docs\apidocs\ru\mephi\dr\parser\retriever\class-use |
| template.attribute.parameters.parameter | Параметр, передаваемый дополнительно в retrieverClass. Список допустимых параметров для каждого из классов описан в документации к соответствующему классу. |
| …parameter.key | Ключ параметра |
| …parameter.value | Значение параметра |
| template.attribute.validators.validator | Валидатор. Отвечает за проверку значения аттрибута |
| …validator.validatorClass | Класс валидатора. Документация по всем валидаторам – в папке docs\apidocs\ru\mephi\dr\parser\validator\class-use |
| …validator.validatorMessage | Сообщение об ошибке, выдаваемое в случае, если документ не является валидным |
| …validator.parameters.parameter | Параметры, передаваемые валидатору на вход |
| …parameter.key | Ключ параметра |
| …parameter.value | Значение параметра |

# Работа с приложением

Для запуска приложения необходимо запустить файл ui.bat. После запуска программы будет предложено ввести команду.

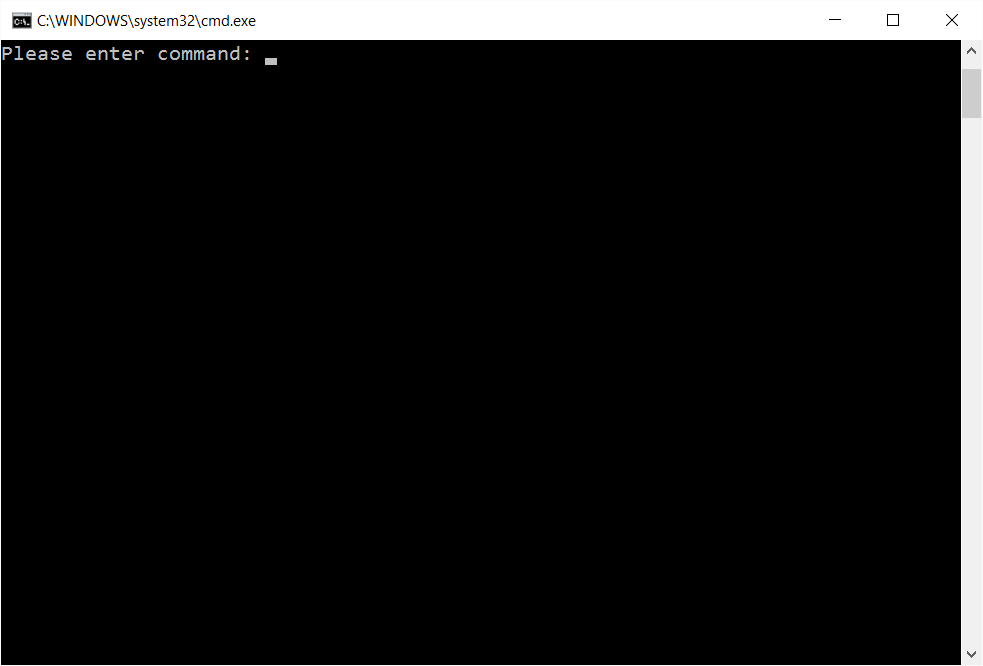


Рисунок . Программа предлагает ввести команду

При неверном вводе команды будет выдана подсказка со списком доступных команд:

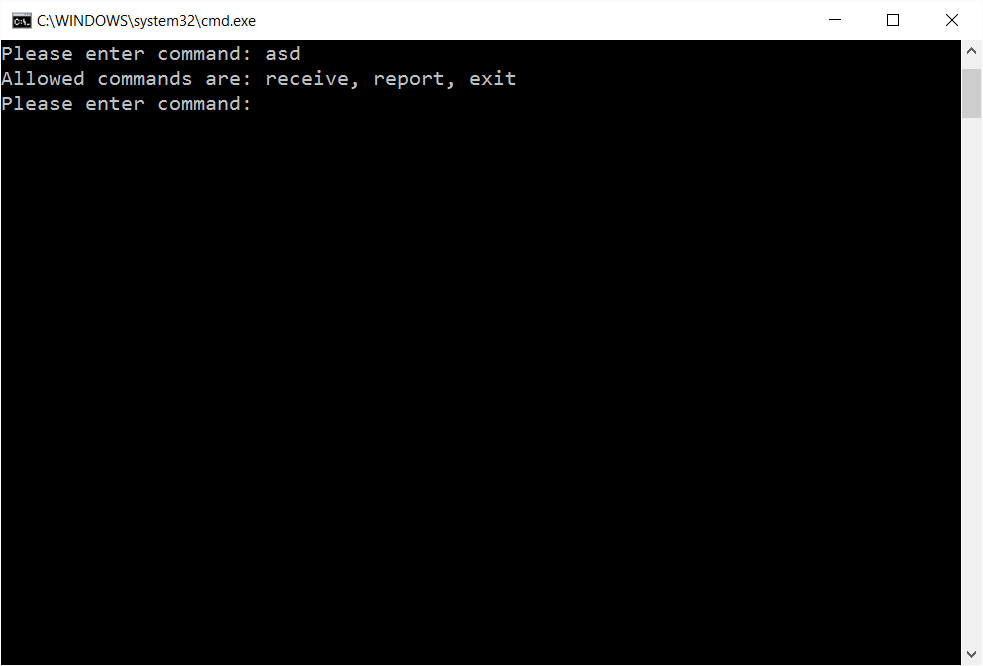


Рисунок . Подсказка - список доступных команд

Список доступных команд:

|  |  |
| --- | --- |
| Команда | Назначение |
| receive | Получить рефераты с электронной почты и отсортировать |
| report | Составление отчетности по всем отсортированным рефератам |
| exit | Выйти из программы |

При выполнении команды receive на экране отобразится логирование операции (включая возможные ошибки):

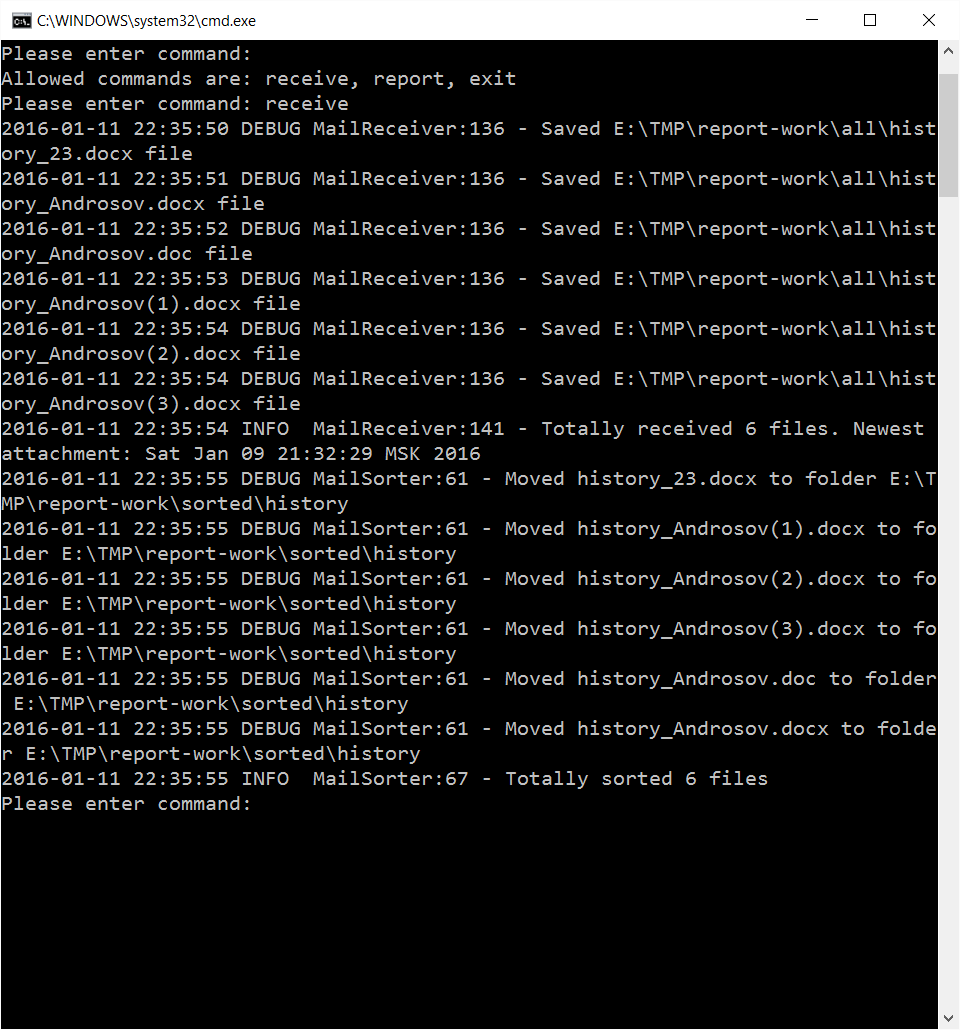


Рисунок . Получение рефератов с почты и сортировка

При вводе команды report программа спросит, отчет с каким расширением необходимо составить:

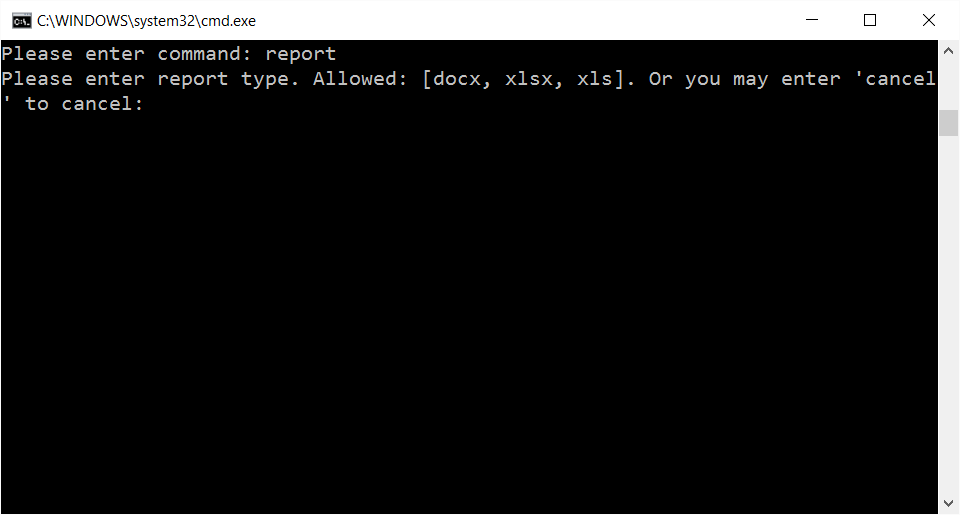


Рисунок . Составление отчета - необходимо выбрать тип отчета

После выбора типа отчета на экране отобразится логирование операции. Скриншот ниже содержит предупреждения о том, что некоторые рефераты не прошли валидацию, т.к. они не являлись рефератами по инстории, но были помещены в папку с рефератами по истории из-за шаблона имени. Консоль Windows настроена неправильно, поэтому кириллица выдается в нечитаемом формате. В конце содержится информационное сообщение о том, что был сохранен отчет, содержащий информацию по 10 рефератам.

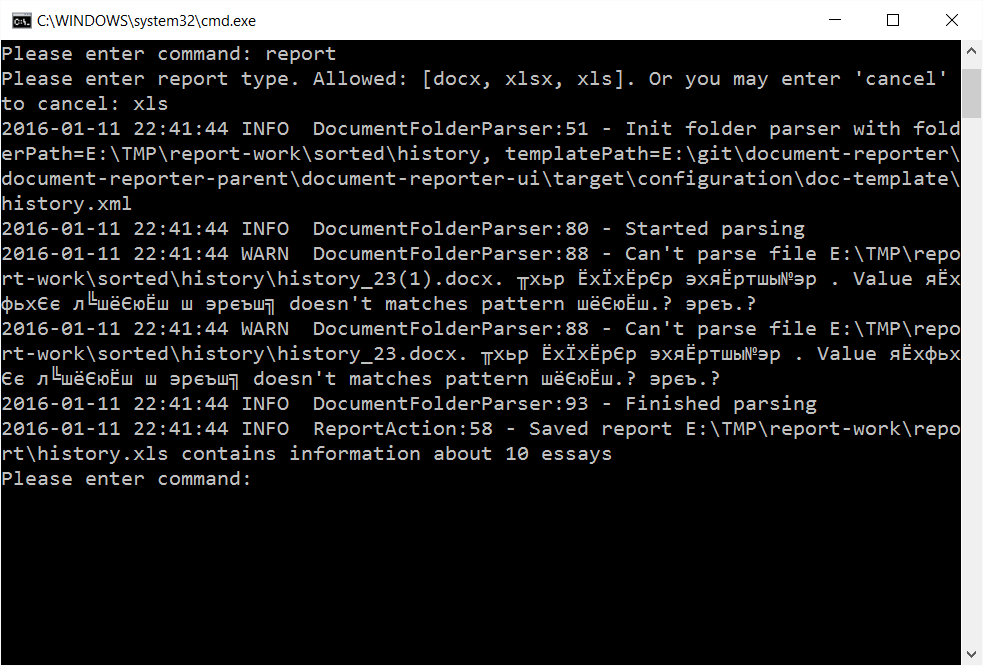


Рисунок . Создание отчета формата xls с предупреждениями