|  |  |
| --- | --- |
|  | «File Converter» Project |
| Requirements  Project Documentation |
|  |  |
| Background | Full set of requirements specification. |
| Purpose | To organize both development and testing process. |
| Scope | Business requirements, user requirements, detailed specification, limitations. |
| Audience | Management staff, project team. |
| File |  |
|  | |

Contents

[1. Project scope 2](#_Toc530682265)

[2. Main goals 3](#_Toc530682266)

[3. Criteria for main goals achievement 3](#_Toc530682267)

[4. Risks 3](#_Toc530682268)

[5. System characteristics 3](#_Toc530682269)

[6. User requirements 3](#_Toc530682270)

[7. Business rules 4](#_Toc530682271)

[8. Quality attributes 4](#_Toc530682272)

[9. Limitations 5](#_Toc530682273)

[10. Detailed specifications 5](#_Toc530682274)

# Project scope

Development of a tool to eliminate encoding multiplicity in text documents stored locally.

# Main goals

* Eliminate the necessity for manual detection and conversion of encoding in text documents.
* Decrease document-processing time by the amount needed for manual encoding detection and conversion.

# Criteria for main goals achievement

* Full automation of encoding detection and conversion.
* Document-processing time reduction by 1-2 minutes (average) due to elimination the necessity for manual encoding detection and conversion.

# Risks

* High complexity of accurate detection of text document initial encoding.

# System characteristics

* SC-1: The application should be a console one.
* SC-2: The application should be developed using PHP (see [L-1](#L_1) for the explanation; PHP-related details are described in [DS-1](#DS_1)).
* SC-3: The application should be a multi-platform one (taking into account [L-4](#L_4)).

# User requirements

* See also the use cases diagram below for details.
* UR-1: Start and stop of the application.
  + UR-1.1: The application start should be performed by the following console command: “php converter.phar SOURCE\_DIR DESTINATION\_DIR [LOG\_FILE\_NAME]” (see [DS-2.1](#DS_2_1) for parameters description, see [DS-2.2](#DS_2_2), [DS-2.3](#DS_2_3), and [DS-2.4](#DS_2_4) for error messages on any misconfiguration situation).
  + UR-1.2: The application stop (shutdown) should be performed by applying Ctrl+C to the console window, which holds the running application.
* UR-2: Configuration of the application.
  + UR-2.1: The only configuration available is through command line parameters (see [DS-2](#DS_2)).
  + UR-2.2: Target encoding for text file conversion is UTF8 (see also [L-5](#L_5)).
* UR-3: Application log.
  + UR-3.1: The application should output its log both to the console and to a log-file (see [DS-4](#DS_4)). Log file name should comply with the rules described in [DS-2.1](#DS_2_1).
  + UR-3.2: Log contents and format are described in [DS-4.1](#DS_4_1), the application reaction to log file presence/absence is described in [DS-4.2](#DS_4_2) and [DS-4.3](#DS_4_3) accordingly.



# Business rules

* BR-1: The source directory and the destination directory
  + BR-1.1: The source directory and the destination directory may NOT be the same directory (see also [DS-2.1](#DS_2_1) and [DS-3.2](#DS_3_2)).
  + BR-1.2: The destination directory may NOT by inside the source directory or any its subdirectories (see also [DS-2.1](#DS_2_1) and [DS-3.2](#DS_3_2)).

# Quality attributes

* QA-1: Performance
  + QA-1.1: The application should provide the processing speed of at least 5 MB/sec with the following (or equivalent) hardware: CPU i7, RAM 4 GB, average disc read/write speed 30 MB/sec. See also [L-6](#L_6).
* QA-2: Resilience to input data
  + QA-2.1: See [DS-5.1](#DS_5_1) for the requirements to input file formats.
  + QA-2.2: See [DS-5.2](#DS_5_2) for the requirements to input file size.
  + QA-2.3: See [DS-5.3](#DS_5_3) for the details on application reaction on incorrect input file format.

# Limitations

* L-1: The application should be developed using PHP as the Customer is going to support the application with his own IT-department.
* L-2: See [DS-1](#DS_1) for PHP version and configuration details.
* L-3: PHP setup and configuration process are out of this project scope and therefore are NOT described in any product/project documentation.
* L-4: Multi-platform capabilities of the application are the next: it should work with Windows and Linux assuming that proper PHP version (see [DS-1.1](#DS_1_1)) works there.
* L-5: The target encoding (UTF8) is fixed. There is no option to change it.
* L-6: The [QA-1.1](#QA_1_1) may be violated in case of objective reasons (e.g., system overload, low-performing hardware and so on).

# Detailed specifications

**DS-1: PHP**

DS-1.1: Minimal version – 5.5.

DS-1.2: The mbstring extension should be installed and enabled.

**DS-2: Command line parameters**

DS-2.1: The application receives three command line parameters during the start process:

* SOURCE\_DIR – mandatory parameter, points to the directory with files to be processed;
* DESTINATION\_DIR – mandatory parameter, points to the directory to store converted files (see also [BR-1.1](#BR_1_1) and [BR-1.2](#BR_1_2));
* LOG\_FILE\_NAME – optional parameter, points to the log file (if omitted, “converter.log” file should be created in the same directory where “converter.phar” is located);

DS-2.2: If some mandatory command line parameter is omitted, the application should shut down displaying standard usage-message (see [DS-3.1](#DS_3_1)).

DS-2.3: If more than three command line parameters are passed to the application, it should ignore any parameter except listed in [DS-2.1](#DS_2_1).

DS-2.4: If the value of any command line parameter is incorrect, the application should shut down displaying standard usage-message (see [DS-3.1](#DS_3_1)) and incorrect parameter name, value, and proper error message (see [DS-3.2](#DS_3_2)).

**DS-3: Messages**

DS-3.1: Usage message: “USAGE converter.phar SOURCE\_DIR DESTINATION\_DIR LOG\_FILE\_NAME”.

DS-3.2: Error messages:

* Directory not exists or inaccessible.
* Destination dir may not reside within source dir tree.
* Wrong file name or inaccessible path.

**DS-4: log**

DS-4.1: The log format is the same for the console and the log file: YYYY-MM-DD HH:II:SS operation\_name operation\_parameters operation\_result.

DS-4.2: If the log-file is missing, a new empty one should be created

DS-4.3: If the log file exists, the application should append new records to its tail.

**DS-5: File format and size**

DS-5.1: The application should process input files in Russian and English languages with the following encodings: WIN1251, CP866, and KOI8R.

Supported file formats (defined by extension) are:

* Plain Text (TXT).
* Hyper Text Markup Language Document (HTML).
* Mark Down Document (MD).

DS-5.2: The application should process files up to 50 MB (inclusive), the application should ignore any file with the size larger than 50 MB.

DS-5.3: If a file with supported format (see [DS-5.1](#DS_5_1)) contains format-incompatible data, such data may be damaged during file processing, and this situation should be treated as correct application work.