Учреждение образования

«БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ»

Кафедра информатики

Отчет по лабораторной работе №3

Доработка требований

Вариант 4

Студент: гр. 853501

Труханович Максим

Руководитель: Позняк В. В.

Минск 2021

**1. Суть проекта**

Разработка инструмента, позволяющего автоматически распознавать кодировку текстовых файлов, расположенных в локальном дисковом хранилище, и преобразовывать данные файлы к выбранной кодировке.

**2. Основные цели**

· Исключение необходимости ручного подбора кодировок текстовых документов.

· Преобразование данного текстового документа к заданной кодировке.

· Сокращение времени работы с текстовым документом на величину, необходимую для ручного подбора кодировки.

**3. Метрики достижения основных целей**

· Полная автоматизация определения и преобразования кодировки текстового документа к заданной.

· Сокращение времени обработки текстового документа на время, необходимое для ручного подбора кодировки.

**4. Риски**

· Высокая техническая сложность безошибочного определения исходной кодировки текстового документа.

· Высокая техническая сложность преобразования исходного документа к нужной кодировке. Ошибка преобразования может привести к повреждению данных документа.

**5. Системные характеристики**

· СХ-1: Приложение является консольным. · СХ-2: Для работы приложение использует интерпретатор PHP (см. О-1 за разъяснениями; соответствующие PHP детали описаны в ДС-1). · СХ-3: Приложение является кроссплатформенным (учитывая О-4).

**6. Пользовательские требования**

Также см. диаграмму вариантов использования.

· ПТ-1: Запуск и остановка приложения.

o ПТ-1.1: Запуск приложения производится из консоли командой “php converter.phar SOURCE\_DIR DESTINATION\_DIR [LOG\_FILE\_NAME]” (см. ДС-2.1 для описания параметров, см. ДС-2.2, ДС-2.3, и ДС-2.4 для сообщений об ошибках в любой ситуации неправильной конфигурации).

o ПТ-1.2: Остановка приложения производится выполнением команды Ctrl+C. В консоль выводится сообщение о том, что операция не была завершена: «Работа программы прервана». Файл для записи удаляется. В лог-файл в результат операции записывается это же сообщение «Работа программы прервана».

· ПТ-2: Конфигурирование приложения.

o ПТ-2.1: Конфигурирование приложения сводится к указанию путей в файловой системе (см. ДС-2). o ПТ-2.2: Целевой кодировкой является UTF8 (см. также О-5).

· ПТ-3: Просмотр журнала работы приложения.

o ПТ-3.1: В процессе работы приложение должно выводить журнал своей работы в консоль и лог-файл (см. ДС-4). Имя лог-файла должно соответствовать правилам, описанным в разделе в ДС-2.1 o ПТ-3.2: Содержание и формат журнала описаны в ДС-4.1, реакция приложения на наличие/отсутствие лог-файла описана в ДС-4.2 и ДС-4.3 соответственно.

**7. Бизнес-требования**

· БП-1: Источник и приёмник файлов

o БТ-1.1: Каталоги, являющиеся источником исходных и приёмником конечных файлов не должны совпадать. (см. также ДС-2.1 и ДС-3.2).

o БТ-1.2: Каталог, являющийся приёмником конечных файлов, не может быть подкаталогом источника. (см. также ДС-2.1 и ДС-3.2)

o БТ-1.3: Каталог, являющийся приёмником конечных файлов, не может совпадать с каталогом лог-файла

**8. Атрибуты качества**

· АК-1: Производительность

o АК-1.1: Приложение должно обеспечивать скорость обработки данных 5 МБ/сек со следующим (или эквивалентным оборудованием): CPU i7, RAM 4 GB, средняя скорость чтения/записи диска 30 МБ/сек. См. также О-6.

· АК-2: Устойчивость к входным данным

o АК-2.1: Требования относительно форматов обрабатываемых файлов изложены в ДС-5.1.

o АК-2.2: Требования относительно размеров обрабатываемых файлов изложены в ДС-5.2.

o АК-2.3: Поведение приложения в ситуации обработки файлов с нарушениями формата определено в ДС-5.3.

**9. Ограничения**

· О-1: Приложение разрабатывается на языке программирования PHP, использование которого обусловлено возможностью заказчика осуществлять поддержку приложения силами собственного IT-отдела.

· О-2: Ограничения относительно версии и настроек интерпретатора PHP отражены в ДС-1.

· О-3: Процедуры установки и настройки интерпретатора PHP выходят за рамки данного проекта и не описываются в документации.

· О-4: Кроссплатформенные возможности приложения сводятся к способности работать под ОС Windows 7, 8, 10 и Linux Debian, Fedora, Astra, Mint, Ubuntu, Manjaro, Arch, Kubuntu, поддержбивающих работу интерпретатора PHP версии (см. ДС-1.1)

· О-5: Целевая кодировка UTF8 является жёстко заданной, и её изменение в процессе эксплуатации приложения не предусмотрено. · О-6: Допускается невыполнение АК-1.1 в случае, если невозможность обеспечить заявленную производительность обусловлена объективными внешними причинами (например, техническими проблемами на сервере заказчика).

**10. Детальные спецификации**

• **ДС-1: Интерпретатор PHP**

ДС-1.1: Минимальная версия – 5.5.

ДС-1.2: Для работы приложения должно быть установлено и включено расширение mbstring для интерпретатора PHP.

• **ДС-2: Параметры командной строки**

ДС-2.1: При запуске приложения оно получает из командной строки три параметра:

• SOURCE\_DIR – обязательный параметр, определяет путь к каталогу с файлами, которые необходимо обработать;

• DESTINATION\_DIR – обязательный параметр, определяет путь к каталогу, в который необходимо поместить обработанные файлы (см. также БТ-1.1 и БТ-1.2);

• LOG\_FILE\_NAME –необязательный параметр, определяет полное имя лог-файла (по умолчанию лог-файл с именем «converter.log» размещается по тому же пути, по которому находится файл скрипта con-verter.php);

ДС-2.2: При указании недостаточного количества параметров командной строки приложение должно завершить работу, выдав сообщение об использовании (см. ДС-3.1).

ДС-2.3: При указании излишнего количества параметров командной строки приложение должно игнорировать все параметры командной строки, кроме указанных в пункте ДС-2.1.

ДС-2.4: При указании неверного значения любого из параметров командной строки приложение должно завершить работу, выдав сообщение об использовании (см. ДС-3.1) а также сообщив имя неверно указанного параметра, его значение и суть ошибки (см. ДС-3.2).

• **ДС-3: Сообщения**

ДС-3.1: Сообщения об использовании: “USAGE converter.phar SOURCE\_DIR DESTINATION\_DIR LOG\_FILE\_NAME”.

ДС-3.2: Сообщения об ошибках:

• Каталог не существует или недоступен.

• Dir назначения может не находиться в дереве dir источника.

• Неправильное имя файла или недоступный путь.

• Расширение файла не соответствует предложенным

• Данные файла повреждены

• **ДС-4: Журнал работы**

ДС-4.1: Формат журнала работы одинаков для отображения в консоли и записи в лог-файл: YYYY-MM-DD HH:II:SS имя\_операции параметры\_операции результат\_операции.

ДС-4.2: В случае если лог-файл отсутствует, должен быть создан новый пустой лог-файл.

ДС-4.3: В случае если лог-файл уже существует, должно происходить добавление новых записей в его конец.

• **ДС-5: Форматы и размеры файлов**

ДС-5.1: Приложение должно обрабатывать текстовые файлы на русском и английском языках в следующих исходных кодировках: WIN1251, CP866, and KOI8R.

Обрабатываемые файлы могут быть представлены в следующих форматах, определяемых расширениями файлов:

• Plain Text (TXT).

• Hyper Text Markup Language Document (HTML).

• Mark Down Document (MD).

ДС-5.2: Приложение должно обрабатывать файлы размером до 50 МБ (включительно), игнорируя любой файл, размер которого превышает 50 МБ. ДС-5.3: Если файл с расширением из ДС-5.1 содержит внутри себя данные, не соответствующие формату файла, или формат файла не соответствует списку предложенных, или кодировка файла не распознается, то допускается повреждение таких данных. В лог-файл записываются в прежнем формате, но в результат операции записы