

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

ОТЧЕТ
по производственной практике
Тема: SlideChecker - критерии проверки ВКР (литература)

Студентка гр. 9382

Сорочина М.В.

Руководитель

Заславский М.М.

Санкт-Петербург

2022

ЗАДАНИЕ НА ПРОИЗВОДСТВЕННУЮ ПРАКТИКУ

Студентка Сорочина М.В.

Группа 9382

Тема практики: SlideChecker - критерии проверки ВКР (литература)

Задание на практику:

Добавить несколько заданных критериев проверки в SlideChecker.

Сроки прохождения практики: 29.06.2022 – 12.07.2022

Дата сдачи отчета: 10.07.2022

Дата защиты отчета: 10.07.2022

Студентка		Сорочина М.В.
Руководитель		Заславский М.М,

АННОТАЦИЯ

Целью прохождения практики является получение навыка участия в проекте, приближенном к коммерческой разработке. Необходимо решить три поставленные задачи, нацеленные на доработку имеющегося продукта. Продуктом является веб-приложение для автоматизированной проверки дипломных презентаций и выпускных квалификационных работ. В результате выполнения задач должны быть добавлены критерии проверки раздела списка источников.

SUMMARY

The purpose of the internship is to gain the skill of participating in a project that is close to commercial development. It is necessary to solve three tasks, aimed at finalizing the existing product. The product is a web application for automated verification of graduation presentations and graduation theses. As a result of completing the tasks, criteria for checking the section of the list of references should be added.

СОДЕРЖАНИЕ

	Введение	5
1.	Первое задание	6
1.1.	Формулировка задачи.	6
1.2.	Выполнение задачи.	6
2.	Второе задание	7
2.1.	Формулировка задачи.	7
2.2.	Выполнение задачи.	7
3.	Третье задание	8
3.1.	Формулировка задачи.	8
3.2.	Выполнение задачи.	8
	Заключение	9
	Использованные технологии.	10
	Приложение А. Отзыв руководителя.	11

ВВЕДЕНИЕ

SlideChecker – веб-приложение для проверки дипломных презентаций и ВКР по заданным критериям. Основные технологии проекта: Python, Flask, Docker.

Задачи практики:

1. Добавить критерий проверки объема текста в страницах.
2. Добавить проверку наличия запретных слов в списке источников.
3. Добавить проверку наличия ссылок на все элементы списка источников.

1. ПЕРВОЕ ЗАДАНИЕ

1.1. Формулировка задачи.

Нужно вычислить объем текста в страницах без разделов «Приложение ...» и сравнить с параметром критерия

1.2. Выполнение задачи.

Для выполнения задачи было изучено содержание репозитория, в особенности классы для обработки документов формата docx и pdf, а также конвертер из docx в pdf и класс, содержащий проверку на пустоту файла. На основании полученных данных, был сделан вывод, что посчитать количество страниц можно при помощи поиска страницы с первым приложением в pdf файле. Впоследствии это решение было реализовано на языке Python и проверка добавлена в систему.

2. ВТОРОЕ ЗАДАНИЕ

2.1. Формулировка задачи.

Проверка на наличие источников с запретными словами, вывод списка не прошедших проверку источников.

Параметр критерия - запретные слова для упоминания в источниках.

2.2. Выполнение задачи.

Был найден список источников при помощи регулярных выражений и создан массив с источниками. В случае не обнаружения списка источников проверка не пройдена и выводится “Нет списка литературы”. Если список найден, то каждое слово приводится в нормальную форму и сравнивается с нормальной формой запрещенных слов. При совпадении в словарь добавляется слово с ключом равным номеру источника. Приведение в нормальную форму осуществляется при помощи модуля `ru morphology2`. В случае, когда ни одного запрещенного слова не было найдено, проверка считается пройденной. Иначе проверка не пройдена и выводится сообщение со списком номеров источников и запрещенными словами в них.

3. ТРЕТЬЕ ЗАДАНИЕ

3.1. Формулировка задачи.

Каждый элемент списка литературы должен иметь ссылку в тексте. виды ссылок: [1] , [2-3] [4,6, 7-8], [5,8].

Критерий проверяет и выводит список проблемных источников.

3.2. Выполнение задачи.

Задача логически разбилась на две основные: поиск количества источников и поиск ссылок на источники в тексте, а затем сравнение полученных данных.

Поиск количества источников осуществлялся при помощи pdf версии файла, так как при обработке docx в Python не сохраняется нумерация списков, в отличие от pdf.

Поиск ссылок на источники был реализован с использованием регулярных выражений. При нахождении очередной ссылки все упомянутые в ней источники добавлялись в множество.

После определения количества источников и нахождения всех ссылок, сравнивались два множества - найденные номера и номера из списка источников. Проверка пройдена, если множества совпали. В случае отсутствия ссылок на один или более источников или наличия ссылок на несуществующие источники выводились соответствующие сообщения.

ЗАКЛЮЧЕНИЕ

В ходе практики были выполнены все поставленные задачи, получен навык работы над проектом в команде. А также получены навыки работы с Docker и такими модулями языка Python, как `pymorphy2` и `python-docx`.

ИСПОЛЬЗОВАННЫЕ ТЕХНОЛОГИИ.

Приложение написано с использованием Python, Flask, Docker.

Все критерии были реализованы на ЯП Python. В основном использовались модули re, python-docx, pdfplumber.

ПРИЛОЖЕНИЕ А.

ОТЗЫВ РУКОВОДИТЕЛЯ.

Отзыв

о прохождении производственной практики

Сорочина Мария Витальевна, студентка группы 9382, третьего курса бакалавриата Санкт-Петербургского электротехнического университета “ЛЭТИ” им В.И. Ульянова, проходил(а) производственную практику по теме “SlideChecker - критерии проверки ВКР (литература)” с 29.06.2022 по 12.07.2022.

В течении практики обучающаяся Сорочина Мария Витальевна выполняла задачи по реализации автоматизированных проверок отчетов.

В результате практики обучающаяся реализовала следующие проверки, успешно интегрированные в систему проверки:

- проверка общего объема документа
- проверка качества списка литературы
- проверка наличия ссылок на список литературы

Получаемую работу обучающаяся Сорочина Мария Витальевна выполнила в полном объеме, самостоятельно решала возникающие трудности, своевременно исправляла ошибки.

По результатам работы Сорочиной Марии Витальевны производственную практику оцениваю на “отлично”.

Руководитель

Заславский М.М