

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

ОТЧЕТ
по производственной практике
Тема: Document Insight System. Новые критерии для ВКР.

Студентка гр. 1384

Усачева Д.В.

Руководитель

Иванов Д.В.

Санкт-Петербург

2024

ЗАДАНИЕ НА ПРОИЗВОДСТВЕННУЮ ПРАКТИКУ

Студентка Усачева Дарья Владимировна

Группа 1384

Тема практики: Document Insight System. Новые критерии для ВКР.

Задание на практику:

Добавить в систему анализа документов новые критерии для проверки отчетов и презентаций ВКР. Список необходимых для создания критериев:

1. Критерий для презентаций – проверка объема текста на каждом слайде;
2. Критерий для презентаций – проверка объема текста для заключения;
3. Критерий для презентаций/отчетов – не пропускать задачи из серии "доделать, решить ...";
4. Критерий для презентаций – проверка наличия обзора в числе задач;
5. Критерий для презентаций – проверка заголовков на дублируемость;
6. Критерий для отчетов – проверка количества абзацев в разделах;
7. Критерий для отчетов – проверка верных значений таблиц титульного листа.

Сроки прохождения практики: 26.06.2024 - 09.07.2024

Дата сдачи отчета: 09.07.2024

Дата защиты отчета: 09.07.2024

Студентка

Усачева Д.В.

Руководитель

Иванов Д.В.

АННОТАЦИЯ

Цель практики: развитие навыков парсинга документов и повышение качества автоматической проверки отчетов и презентаций ВКР за счет создания новых критериев, таких как:

1. Критерий для презентаций – проверка объема текста на каждом слайде;
2. Критерий для презентаций – проверка объема текста для заключения;
3. Критерий для презентаций/отчетов – не пропускать задачи из серии "доделать, решить ...";
4. Критерий для презентаций – проверка наличия обзора в числе задач;
5. Критерий для презентаций – проверка на дублируемость заголовков слайдов;
6. Критерий для отчетов – проверка количества абзацев в разделах;
7. Критерий для отчетов – проверка верных значений таблиц титульного листа.

SUMMARY

The practice goal is to develop document parsing skills and enhance the quality of automatic verification of final qualification work (FQW) reports and presentations by creating new criteria such as:

1. Criterion for presentations – checking the amount of text on each slide;
2. Criterion for presentations – checking the amount of text for the conclusion;
3. Criterion for presentations/reports– not to overlook tasks like "to finish, to solve ...";
4. Criterion for presentations – checking for a review among the tasks;
5. Criterion for presentations – checking for duplication of slide titles;
6. Criterion for reports – checking the number of paragraphs in sections;
7. Criterion for reports – verifying the correct values in the title page tables.

СОДЕРЖАНИЕ

	Введение	5
1.	Результаты работы	6
1.1.	Выполненные задачи	6
2.	Использованные технологии	7
2.1.	Список технологий	7
	Заключение	8

ВВЕДЕНИЕ

Автоматическая проверка ВКР необходима, поскольку она обеспечивает более высокую степень объективности и согласованности при оценке научных работ. Это позволяет избежать человеческой ошибки и предвзятости, а также значительно ускоряет процесс базовой проверки большого количества документов. Кроме того, автоматизация способствует стандартизации критериев оценки, что важно для обеспечения равных условий для всех студентов.

Целью практики является совершенствование умений анализа документов и повышение эффективности автоматической проверки отчетов и презентаций ВКР путем введения новых параметров оценки. Эти параметры включают анализ количества текста на слайдах презентаций, оценку объема заключения, исключение неправильно формулируемых задач, проверку на наличие обзора среди поставленных задач, контроль за уникальностью заголовков слайдов, а также оценку количества абзацев в разделах отчетов и корректность данных в таблицах титульного листа.

1. РЕЗУЛЬТАТ РАБОТЫ

1.1. Выполненные задачи

1. Перед тем как приступить к разработке новых критериев был проведен анализ проекта «Document Insight System». Для этого необходимо было изучить его код и документацию.

2. Разработан и впоследствии улучшен класс `SlideTextVolumeCheck` для проверки соответствия объема текста на каждом слайде заданным критериям. Улучшение представляет собой добавление режима работы для анализ конкретного слайда (Заключения).

3. Улучшен класс `ReportMainCharacterCheck`, в котором ранее проводилась проверка только фамилии и должности заведующего кафедрой. Теперь данный класс осуществляет проверку составляющих таблиц титульного листа, задания и календарного плана.

4. Разработан класс `SlideHeadersDuplicationCheck` для проверки заголовков слайдов на дублируемость.

5. Разработаны классы `TaskTracker` и `ReportTaskTracker` для поиска недопустимых слов в формулировках задач в презентациях и отчетах соответственно.

6. Разработан класс `OverviewInTasks` для поиска обязательных слов в формулировках задач в презентациях.

7. Разработан класс `ReportParagraphsCountCheck` для проверки соответствия количества абзацев в главах и их подразделах заданным критериям.

2. ИСПОЛЬЗОВАННЫЕ ТЕХНОЛОГИИ

2.1. Список технологий

Для реализации задачи использовались следующие технологии:

- Python
- Docker

ЗАКЛЮЧЕНИЕ

По итогам производственной практики результаты работы полностью соответствуют поставленной цели. Были разработаны и улучшены инструменты для анализа документов, что позволяет повысить качество автоматической проверки отчетов и презентаций ВКР. Повышены навыки парсинга документов. Введение новых параметров оценки делает процесс проверки более всесторонним и точным.