ТЕХНИЧЕСКОЕ ЗАДАНИЕ

на разработку программного продукта

"Инструмент обработки неструктурированных данных"

Студент Веремчук Илья

ОГЛАВЛЕНИЕ

1. ОБЩИЕ СВЕДЕНИЯ……………………………………………………..….....3

2. ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ……………………………………………...3

3. ТРЕБОВАНИЯ К ДОКУМЕНТАЦИИ………………………………………...4

4. СТАДИИ И ЭТАПЫ РАЗРАБОТКИ……………………………………………5

5. ПОРЯДОК КОНТРОЛЯ И ПРИЕМКИ………………………………………..5

6. ТРЕБОВАНИЯ К ПАТЕНТНОЙ ЧИСТОТЕ…………………………………..5

7. ТРЕБОВАНИЯ ПО СТАНДАРТИЗАЦИИ И УНИФИКАЦИИ……………..6

8. ТРЕБОВАНИЯ К БЕЗОПАСНОСТИ………………………………………….6

1. ОБЩИЕ СВЕДЕНИЯ

1.1. Наименование продукции

Полное наименование: Программный комплекс обработки неструктурированных данных

Краткое наименование: ПКОНД

1.2. Назначение продукции

Программный продукт предназначен для автоматизированной обработки, анализа и структурирования неструктурированных текстовых данных.

1.3. Основание для разработки

- Задание на выпускную квалификационную работу

- Потребность в инструментах обработки неструктурированных данных

2. ТЕХНИЧЕСКИЕ ТРЕБОВАНИЯ

2.1. Требования к функциональным характеристикам

2.1.1. Программный продукт должен обеспечивать:

- Загрузку неструктурированных текстовых данных различных форматов (txt, doc, docx, pdf)

- Предварительную обработку текста (очистка, токенизация)

- Извлечение ключевых слов и фраз

- Классификацию текстов по категориям

- Кластеризацию документов

- Построение семантических связей

- Экспорт результатов в структурированном виде

2.1.2. Требования к надежности:

- Время восстановления после отказа - не более 5 минут

- Среднее время наработки на отказ - не менее 1000 часов

- Сохранение данных при аварийном завершении работы

2.2. Требования к техническому обеспечению

2.2.1. Минимальные системные требования:

- Процессор: Intel Core i5 или аналог

- ОЗУ: 8 ГБ

- Свободное место на диске: 10 ГБ

- ОС: Windows 10/11 или Linux Ubuntu 20.04 и выше

2.3. Требования к программному обеспечению

- Python 3.8 или выше

- Необходимые библиотеки: NLTK, scikit-learn, pandas, numpy

- База данных: PostgreSQL 12 или выше

2.4. Требования к интерфейсу

- Графический пользовательский интерфейс

- Интуитивно понятный дизайн

- Наличие справочной системы

- Поддержка русского и английского языков

3. ТРЕБОВАНИЯ К ДОКУМЕНТАЦИИ

3.1. Состав документации:

- Руководство пользователя

- Руководство администратора

- Программа и методика испытаний

- Техническая документация

4. СТАДИИ И ЭТАПЫ РАЗРАБОТКИ

4.1. Техническое задание - 2 недели

4.2. Эскизный проект - 3 недели

4.3. Технический проект - 4 недели

4.4. Рабочий проект - 8 недель

4.5. Внедрение - 3 недели

5. ПОРЯДОК КОНТРОЛЯ И ПРИЕМКИ

5.1. Виды испытаний:

- Предварительные испытания

- Опытная эксплуатация

- Приемочные испытания

5.2. Общие требования к приемке работы:

- Соответствие требованиям технического задания

- Успешное прохождение всех видов испытаний

- Наличие полного комплекта документации

6. ТРЕБОВАНИЯ К ПАТЕНТНОЙ ЧИСТОТЕ

6.1. Проверка патентной чистоты должна быть выполнена в отношении РФ

7. ТРЕБОВАНИЯ ПО СТАНДАРТИЗАЦИИ И УНИФИКАЦИИ

7.1. Разработка должна вестись с учетом действующих стандартов:

- ГОСТ Р ИСО/МЭК 12207-2010

- ГОСТ 34.601-90

- ГОСТ Р 51904-2002

8. ТРЕБОВАНИЯ К БЕЗОПАСНОСТИ

8.1. Программный продукт должен обеспечивать:

- Защиту от несанкционированного доступа

- Резервное копирование данных

- Протоколирование действий пользователей