

**Федеральное государственное автономное образовательное учреждение  
высшего образования  
«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**Факультет информационных технологий  
Кафедра «Информатика и вычислительная техника»**

**Направление подготовки/ специальность: системная и программная инженерия**

**ОТЧЕТ**  
**по проектной практике**

Студент: Ланщиков Егор Алексеевич      Группа: 241-326

Место прохождения практики: Московский Политех, кафедра «Информатика  
и вычислительная техника»

Отчет принят с оценкой \_\_\_\_\_ Дата \_\_\_\_\_

Руководитель практики: Чернова Вера Михайловна, кафедра "Информатика  
и вычислительная техника"

Москва 2025

## СОДЕРЖАНИЕ

|  |           |
|--|-----------|
| <b>ВВЕДЕНИЕ .....</b>                                      | <b>3</b>  |
| <b>СВЕДЕНИЯ О ПРОЕКТЕ .....</b>                            | <b>4</b>  |
| <b>1.1 Наименование проекта .....</b>                      | <b>4</b>  |
| <b>1.2 Цели и задачи проекта .....</b>                     | <b>4</b>  |
| <b>2.2 Внешний вид и функции: .....</b>                    | <b>5</b>  |
| <b>ЛИЧНЫЙ ВКЛАД В ПРОЕКТ .....</b>                         | <b>7</b>  |
| <b>1. Разработка интерфейса .....</b>                      | <b>7</b>  |
| <b>2. Создание основной документации для проекта .....</b> | <b>11</b> |
| <b>ОБЩАЯ ХАРАКТЕРИСТИКА ОРГАНИЗАЦИИ .....</b>              | <b>13</b> |
| <b>2.1 Наименование заказчика .....</b>                    | <b>13</b> |
| <b>2.3 Деятельность организации .....</b>                  | <b>13</b> |
| <b>ЗАДАНИЕ ПО ПРОЕКТНОЙ ПРАКТИКЕ .....</b>                 | <b>14</b> |
| <b>Настройка Git и репозитория .....</b>                   | <b>14</b> |
| <b>Документирование в формате Markdown .....</b>           | <b>14</b> |
| <b>Разработка статического веб-сайта .....</b>             | <b>14</b> |
| <b>Взаимодействие с организацией-партнёром .....</b>       | <b>15</b> |
| <b>Подготовка итогового отчета .....</b>                   | <b>15</b> |
| <b>ОПИСАНИЕ РЕЗУЛЬТАТОВ ПО ПРОЕКТНОЙ ПРАКТИКЕ .....</b>    | <b>16</b> |
| <b>ЗАКЛЮЧЕНИЕ .....</b>                                    | <b>20</b> |
| <b>СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ .....</b>                | <b>21</b> |

## ВВЕДЕНИЕ

В условиях стремительного развития информационных технологий, когда программное обеспечение и веб-приложения становятся ключевыми элементами различных сфер деятельности, владение современными инструментами и методологиями разработки является неотъемлемым требованием к специалисту в области IT. Данный отчет посвящен рассмотрению проектной практики, направленной на закрепление базовых навыков, формирующих основу для дальнейшего профессионального становления в сфере информационных технологий.

В рамках практики особое внимание уделялось освоению системы контроля версий Git, использованию языка разметки Markdown для подготовки технической документации, а также разработке статических веб-страниц с применением HTML и CSS. Работа с GitHub позволила на практике изучить процессы управления репозиториями, отслеживания изменений и организации совместной деятельности в команде. Освоение Markdown обеспечило формирование компетенций в области структурирования и оформления документации, востребованных в профессиональной среде. Разработка веб-сайтов способствовала практическому закреплению знаний в области верстки и визуального представления информации.

Таким образом, проектная практика стала важным этапом формирования базовых умений в области разработки и сопровождения программных решений, обеспечив сочетание теоретических знаний и практического опыта.

# СВЕДЕНИЯ О ПРОЕКТЕ

## 1.1 Наименование проекта

Полное название: многоплатформенное приложение для анализа текстовых данных с применением технологий обработки естественного языка (NLP) «BindWord XP».

Сокращённое название: «BindWord XP».

## 1.2 Цели и задачи проекта

Цель проекта заключалась в разработке и выпуске кроссплатформенного настольного приложения с графическим интерфейсом, обеспечивающего анализ текстов средствами NLP. Основой послужил Python-скрипт с использованием библиотеки spaCy, адаптированный для интеграции в полноценный программный продукт.

Основные задачи проекта включали:

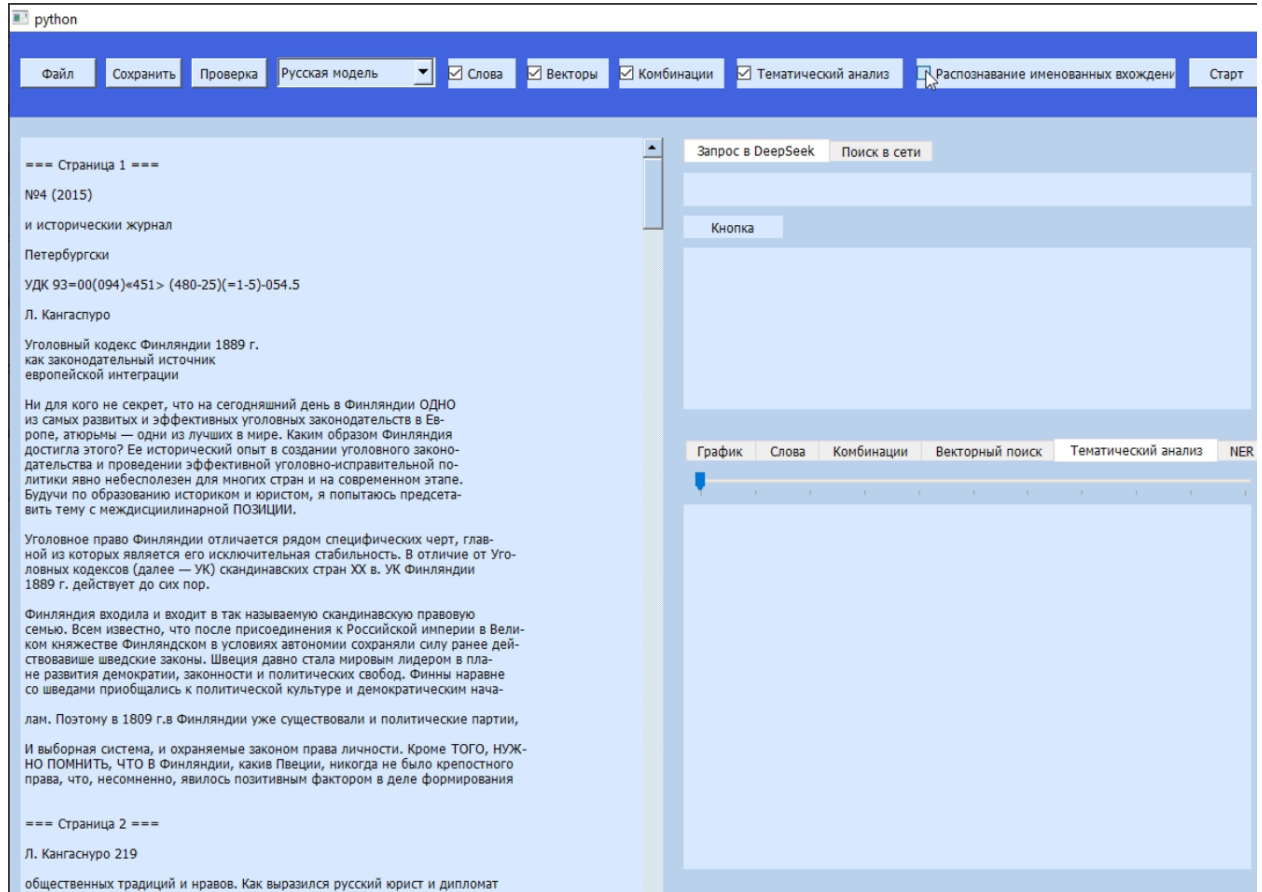
разработку пользовательского интерфейса на базе PyQt;

интеграцию существующего NLP-скрипта в десктопное приложение;

проведение нагрузочного тестирования и оптимизацию производительности;

подготовку пользовательской и технической документации;

## 2.2 Внешний вид и функции:



### 1. Файл

Меню для загрузки и сохранения текстов. Функции: открытие текстового документа, автоматическое помещение его в поле ввода, а также завершение работы приложения.

### 2. Сохранить

Сохраняет результаты анализа в выбранный формат: текст (.txt), таблицу (.csv) или структуру данных (.json).

### 3. Проверка

Кнопка запуска предварительной проверки текста. Может использоваться перед запуском более сложного анализа.

### 4. Слова

Запускает анализ слов: подсчёт общего количества слов, определение самых часто встречающихся, вывод статистики.

## **5. Векторы**

Представляет каждое слово или предложение в виде числового вектора. Это нужно, чтобы программа могла «понимать» смысл текста на математическом уровне. Методы, например: Word2Vec, TF-IDF. Векторы можно использовать для кластеризации, сравнения текстов и визуализации.

## **6. Комбинации**

Анализирует, какие слова часто встречаются рядом друг с другом. Например, «искусственный интеллект» — это биграмма. Программа ищет такие пары и тройки слов (n-граммы), подсчитывает их частотность и показывает, в каком контексте они чаще всего встречаются.

## **7. Тематический анализ**

Позволяет определить, о чём идет речь в тексте. С помощью алгоритмов (например, LDA или NMF) выявляются скрытые темы, которым принадлежат слова. Это полезно для анализа больших объёмов текста — например, новостей или отзывов.

## **8. Распознавание именованных сущностей (NER)**

Определяет и выделяет имена людей, организации, даты, локации и другие важные элементы в тексте. Отображает их отдельно, что удобно для поиска конкретных фактов.

## **9. Переключатель модели (Русская/Английская)**

Позволяет выбрать, на каком языке будет происходить анализ. Для корректной работы с русским и английским текстом используются разные языковые модели.

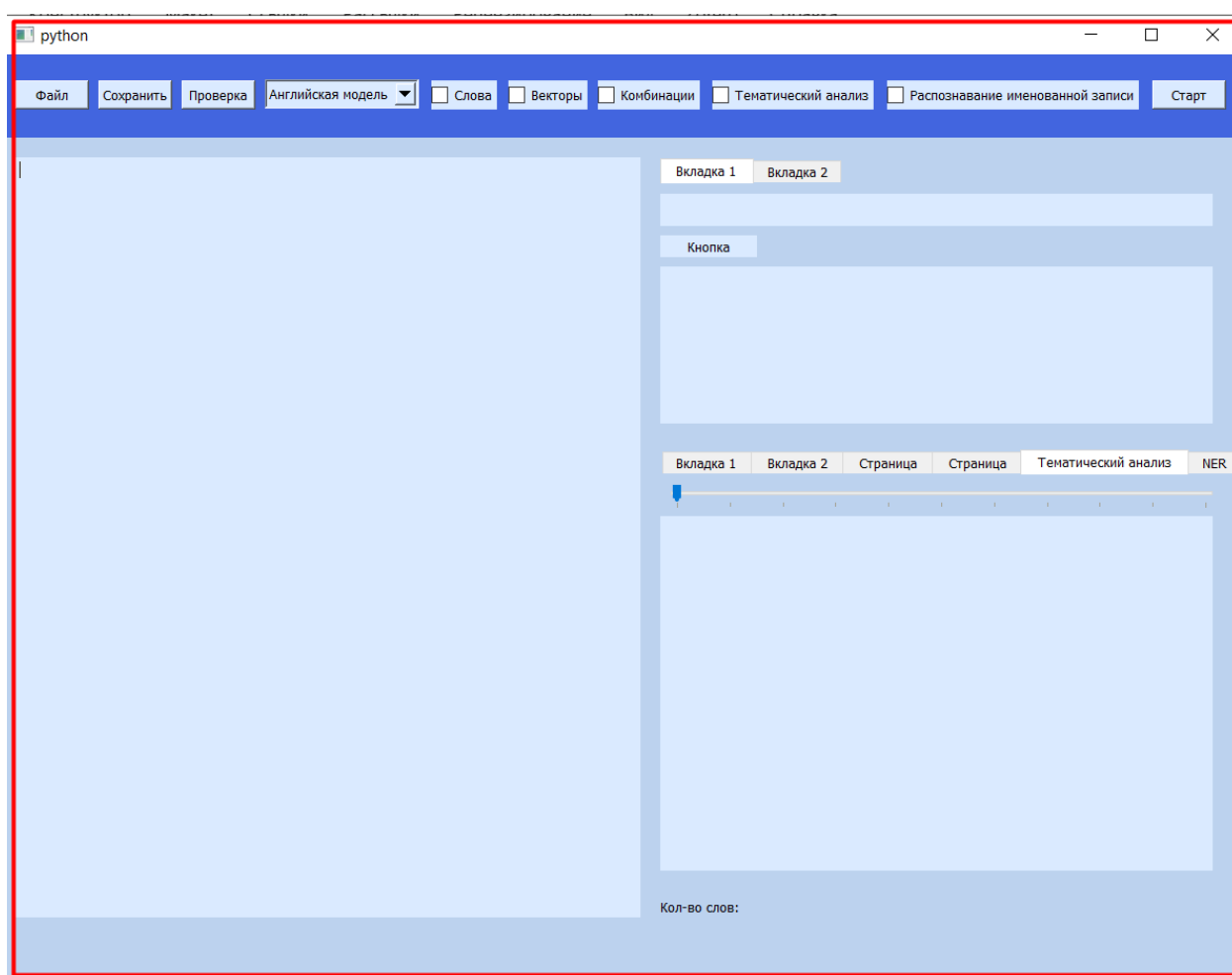
## **10. Кнопка Старт**

Запускает выбранные анализы в зависимости от установленных галочек и загруженного текста.

# ЛИЧНЫЙ ВКЛАД В ПРОЕКТ

## 1. Разработка интерфейса

В рамках индивидуального вклада в проект мной была реализована часть пользовательского интерфейса — главное окно приложения BindWord XP, обеспечивающее структурированное представление функциональных элементов.



Реализованный код для создания основного окна приложения:

```

class Ui_MainWindow(object): 1 usage  ⚡FaddevGleb +1
    def setupUi(self, MainWindow): 1 usage  ⚡FaddevGleb +1
        self.VectorMap = {}
        self.searchEngine = tools.InternetParser.SearchEngine()
        self.ocr = tools.OCR.OCR("D:/ProgramFiles/Tesseract-OCR/tesseract.exe")
        self.en_model = spacy.load("en_core_web_lg")
        self.ru_model = spacy.load("ru_core_news_lg")
        self.models = {"Английская модель": ("en", self.en_model), "Русская модель": ("ru", self.ru_model)}

        MainWindow.setObjectName("MainWindow")
        MainWindow.setEnabled(True)
        MainWindow.resize(1191, 894)
        MainWindow.setStyleSheet("background-color: rgb(188, 210, 238);")
        self.centralwidget = QtWidgets.QWidget(MainWindow)
        self.centralwidget.setObjectName("centralwidget")
        self.tabWidget = QtWidgets.QTabWidget(self.centralwidget)
        self.tabWidget.setGeometry(QtCore.QRect(630, 380, 591, 491))
        self.tabWidget.setStyleSheet("border-color: rgb(218, 234, 255);\n"
                                     "border: 1px;")
        self.tabWidget.setTabPosition(QtWidgets.QTabWidget.North)
        self.tabWidget.setObjectName("tabWidget")
        self.tab = QtWidgets.QWidget()
        self.tab.setObjectName("tab")

```

Текст кода для создания основного окна приложения:

```

def setupUi(self, MainWindow):
    self.VectorMap = {}
    self.searchEngine = tools.InternetParser.SearchEngine()
    self.ocr = tools.OCR.OCR("D:/ProgramFiles/Tesseract-OCR/tesseract.exe")
    self.en_model = spacy.load("en_core_web_lg")
    self.ru_model = spacy.load("ru_core_news_lg")
    self.models = {"Английская модель": ("en", self.en_model),
                  "Русская модель": ("ru", self.ru_model)}

    MainWindow.setObjectName("MainWindow")
    MainWindow.setEnabled(True)
    MainWindow.resize(1191, 894)
    MainWindow.setStyleSheet("background-color: rgb(188, 210, 238);")
    self.centralwidget = QtWidgets.QWidget(MainWindow)
    self.centralwidget.setObjectName("centralwidget")

```

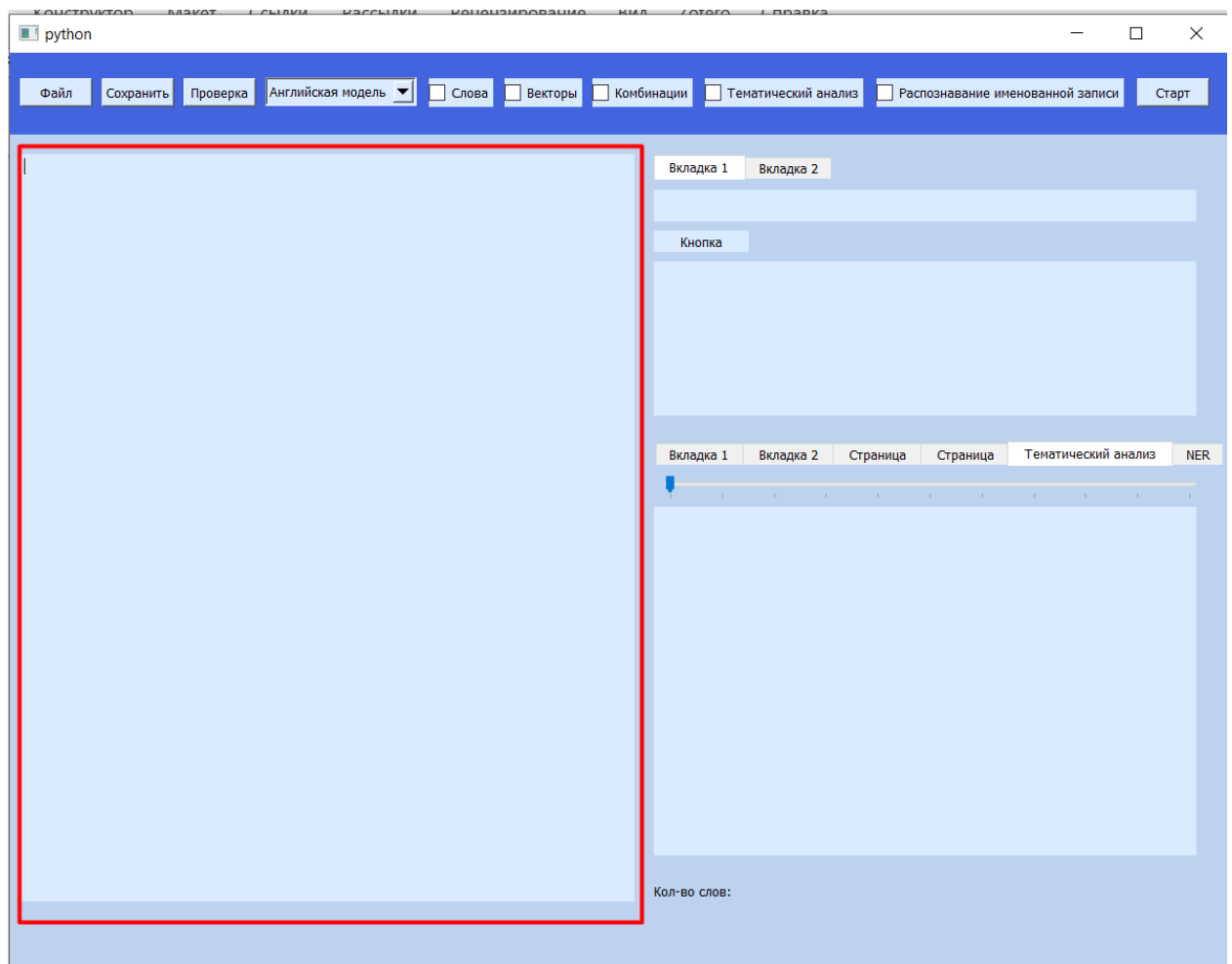


```

self.tab.setObjectName("tab")
self.graphicsView = QtWidgets.QGraphicsView(self.tab)
self.graphicsView.setGeometry(QtCore.QRect(0, 10, 531, 371))
self.graphicsView.setStyleSheet("background-color: rgb(218, 234, 255);")
self.graphicsView.setFrameShape(QtWidgets.QFrame.NoFrame)
self.graphicsView.setObjectName("graphicsView")
self.tabWidget.addTab(self.tab, "")
self.tab_2 = QtWidgets.QWidget()
self.tab_2.setObjectName("tab_2")
self.textBrowser_2 = QtWidgets.QTextBrowser(self.tab_2)
self.textBrowser_2.setGeometry(QtCore.QRect(0, 10, 531, 371))
self.textBrowser_2.setStyleSheet("background-color: rgb(218, 234, 255);\n"
    "")
self.textBrowser_2.setFrameShape(QtWidgets.QFrame.NoFrame)
self.textBrowser_2.setObjectName("textBrowser_2")
self.tabWidget.addTab(self.tab_2, "")
self.tab_3 = QtWidgets.QWidget()
self.tab_3.setObjectName("tab_3")
self.textBrowser_3 = QtWidgets.QTextBrowser(self.tab_3)
self.textBrowser_3.setGeometry(QtCore.QRect(0, 40, 531, 341))
self.textBrowser_3.setStyleSheet("background-color: rgb(218, 234, 255);\n"
    "")

```

Также в числе выполненных мной задач в рамках разработки приложения BindWord XP было создание основного текстового поля на главном экране, обеспечивающего базовую функциональность ввода и редактирования данных пользователем.



Реализованный код для создания поля:

```

self.tab_2.setObjectName("tab_2")
self.textBrowser_2 = QtWidgets.QTextBrowser(self.tab_2)
self.textBrowser_2.setGeometry(QtCore.QRect(0, 10, 531, 371))
self.textBrowser_2.setStyleSheet("background-color: rgb(218, 234, 255);\n"
    "")
self.textBrowser_2.setFrameShape(QtWidgets.QFrame.NoFrame)
self.textBrowser_2.setObjectName("textBrowser_2")
self.tabWidget.addTab(self.tab_2, "")
self.tab_3 = QtWidgets.QWidget()
self.tab_3.setObjectName("tab_3")
self.textBrowser_3 = QtWidgets.QTextBrowser(self.tab_3)
self.textBrowser_3.setGeometry(QtCore.QRect(0, 40, 531, 341))
self.textBrowser_3.setStyleSheet("background-color: rgb(218, 234, 255);\n"
    "")

```

Текст кода:

```

self.textBrowser_2 = QtWidgets.QTextBrowser(self.tab_2)
self.textBrowser_2.setGeometry(QtCore.QRect(0, 10, 531, 371))
self.textBrowser_2.setStyleSheet("background-color:    rgb(218,    234,
255);\n"
    "")
self.textBrowser_2.setFrameShape(QtWidgets.QFrame.NoFrame)
self.textBrowser_2.setObjectName("textBrowser_2")

```

## 2. Создание основной документации для проекта

Мной была подготовлена основная документация к приложению, отражающая ключевые функциональные возможности, структуру интерфейса и принципы взаимодействия пользователя с программой. Также документация описывает архитектуру главного окна, назначение панелей управления и элементов интерфейса, функционал инструментов анализа текстовых данных.

Вид файла (ссылка в используемых источниках):

# Документация к приложению «BindWord XP» для анализа текстовых данных (NLP)

## Основной интерфейс

Главное окно (MainWindow) — это центр работы пользователя с программой. Здесь можно ввести текст, запустить нужный тип анализа и посмотреть результаты в виде таблиц, графиков и текстов.

## Панель управления (верхняя часть экрана)

### 1. Файл

Меню для загрузки и сохранения текстов. Функции: открытие текстового документа, автоматическое помещение его в поле ввода, а также завершение работы приложения.

### 2. Сохранить

Сохраняет результаты анализа в выбранный формат: текст (.txt), таблицу (.csv) или структуру данных (.json).

### 3. Проверка

Кнопка запуска предварительной проверки текста. Может использоваться перед запуском более сложного анализа.

---

## **ОБЩАЯ ХАРАКТЕРИСТИКА ОРГАНИЗАЦИИ**

### **2.1 Наименование заказчика**

Федеральное государственное автономное образовательное учреждение высшего образования «Московский политехнический университет».

### **2.2 Организационная структура**

Ректорат;

Административные подразделения;

Факультеты;

Кафедры.

Научные и исследовательские центры;

Студенческое самоуправление;

### **2.3 Деятельность организации**

Московский политехнический университет — многопрофильное высшее учебное заведение, являющееся участником программы «Приоритет 2030». Учредителем вуза выступает Министерство науки и высшего образования Российской Федерации.

## **ЗАДАНИЕ ПО ПРОЕКТНОЙ ПРАКТИКЕ**

### **Настройка Git и репозитория**

- создание индивидуального репозитория на GitHub
- освоение базовых команд Git (клонирование, фиксация изменений, работа с ветками);
- регулярное ведение истории проекта с использованием содержательных сообщений к коммитам.

### **Документирование в формате Markdown**

- оформление материалов проекта (описание, журнал прогресса, отчетность) в формате Markdown;
- освоение синтаксиса разметки и подготовка необходимых документов.

### **Разработка статического веб-сайта**

- создание сайта о проекте с использованием HTML и CSS;
- уникальность наполнения и дизайна не менее 50 % относительно аналогичных работ;
- включение в структуру сайта следующих разделов:
  1. главная страница с аннотацией;
  2. страница «О проекте»;

3. раздел «Участники» с указанием вклада каждого;
4. раздел «Журнал» с публикацией минимум трёх этапов прогресса;
5. страница «Ресурсы» с подборкой полезных ссылок;
6. оформление с использованием графических и мультимедийных материалов;

Примерные трудозатраты: настройка и изучение — 10–14 часов, дизайн и наполнение — 4–8 часов.

### **Взаимодействие с организацией-партнёром**

участие в мероприятиях, организуемых партнёрами (визит, онлайн-встреча, конференции, митапы, семинары и др.);

отражение полученного опыта и выводов в отчётной документации в формате Markdown;

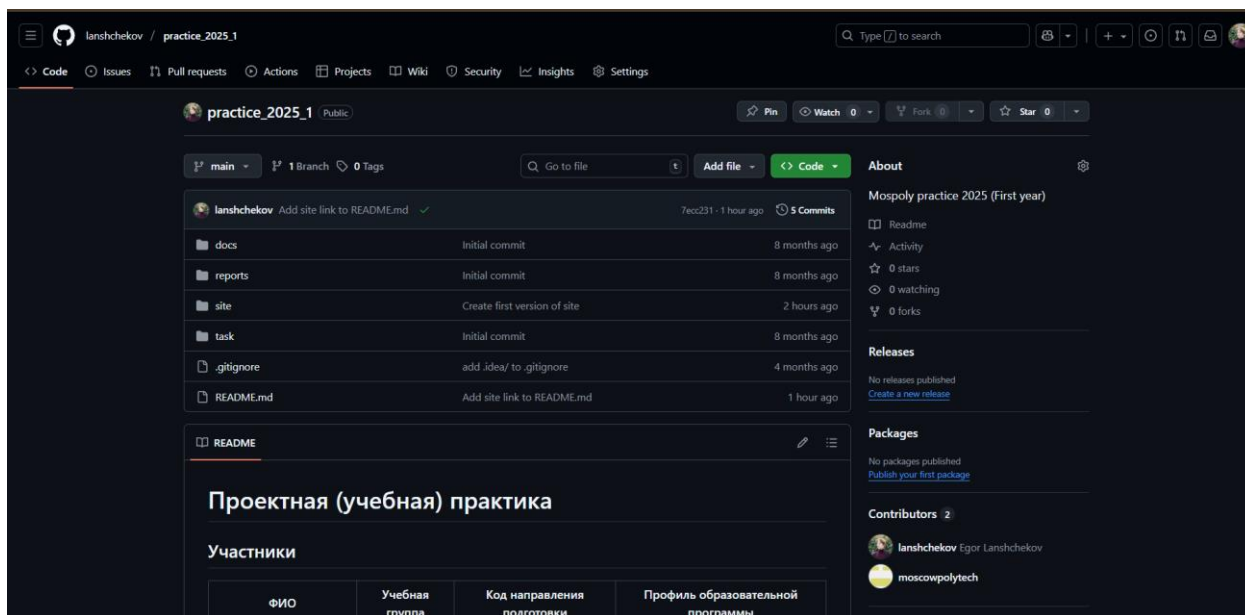
добавление отчета в репозиторий и на сайт проекта.

### **Подготовка итогового отчета**

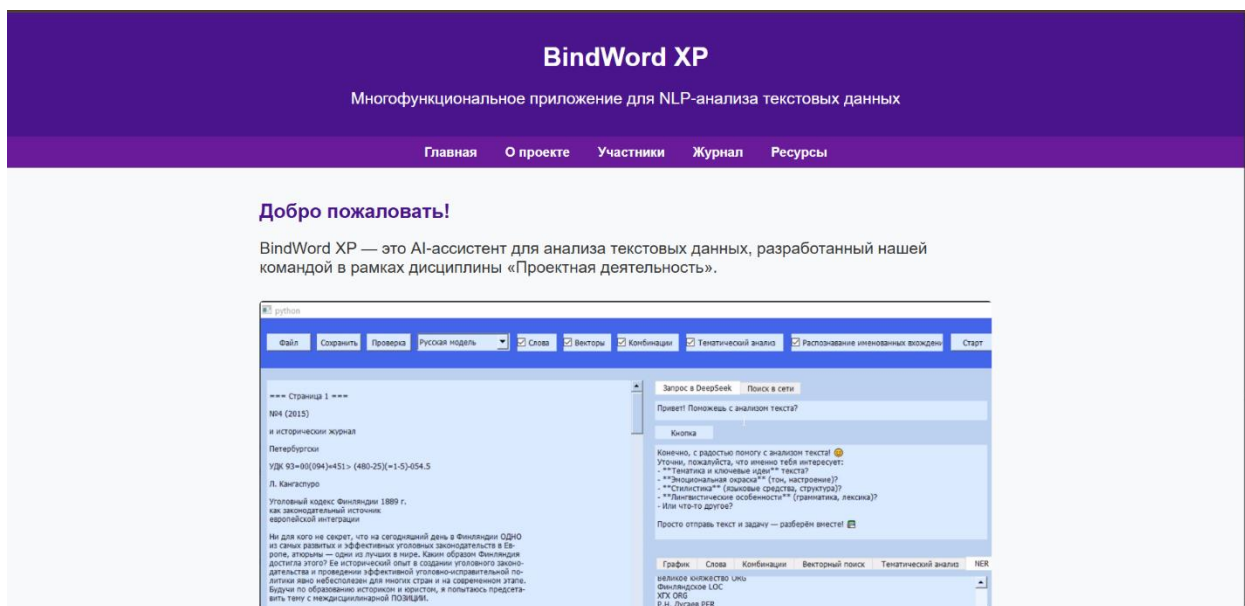
оформление итогового отчета по учебной (проектной) практике в соответствии с установленными требованиями.

# ОПИСАНИЕ РЕЗУЛЬТАТОВ ПО ПРОЕКТНОЙ ПРАКТИКЕ

1. Создан отдельный репозиторий на Github для дальнейшего выполнения заданий



2. Создан статический сайт с описанием основного проекта по дисциплине «Проектная деятельность» под названием «BindWord XP»





О проекте

ГлавнаяО проектеУчастникиЖурналРесурсы

Описание

BindWord XP — это кроссплатформенное приложение для NLP-анализа текстов. Оно позволяет загружать документы в форматах PDF, DOCX, TXT и автоматически выполнять анализ:

- Статистика текста: количество слов, ошибок, ключевые идеи
- Частотный анализ слов и облако тегов
- NER — распознавание именованных сущностей
- Тематический анализ документа
- Экспорт результатов в различные форматы
- Интеграция с AI-помощником DeepSeek

Проект разрабатывается в рамках дисциплины "Проектная деятельность" Московского Политехнического университета.

Куратор проекта: Самелик Юрий Леонидович.

Проблематика проекта

Многие люди при работе с большим объемом информации сталкиваются со снижением продуктивности и риском ошибок из-за ручного анализа, который занимает много времени. Отсутствие программного обеспечения для анализа больших объемов текстовых данных создает значительные трудности для пользователей.

Проблематика проекта

Многие люди при работе с большим объемом информации сталкиваются со снижением продуктивности и риском ошибок из-за ручного анализа, который занимает много времени. Отсутствие программного обеспечения для анализа больших объемов текстовых данных создает значительные трудности для пользователей.

Цель проекта

Разработать и выпустить многоплатформенное программу с графическим интерфейсом для NLP-анализа текстовых данных, основанное на Python/Spacy скрипте, и возможностью обработки текста.

Задачи

- Определение необходимого функционала.
- Написание графического интерфейса.
- Оптимизация кода.
- Дополнение функционала.
- Тестирование и публикация бинарных файлов и исходного кода

Карта Стейкхолдеров

Общественные организации

Медиа

Блоги

Издатели

Московский политех

Научные сообщества, союзы

Карта Стейкхолдеров

Общественные организации

Медиа

Блоги

Издатели

Журналы

Службы СММ университетов

Журналисты

Маркетологи

ВУЗы

Ученые

Приложение с графическим пользовательским интерфейсом для анализа текстовых данных

Московский политех

Научные сообщества, союзы



Страница участников в подгруппе выполнения проектной практики:

**Участники проекта**

Главная   О проекте   Участники   Журнал   Ресурсы

**Ланцев Егор Алексеевич (группа 241-326)**

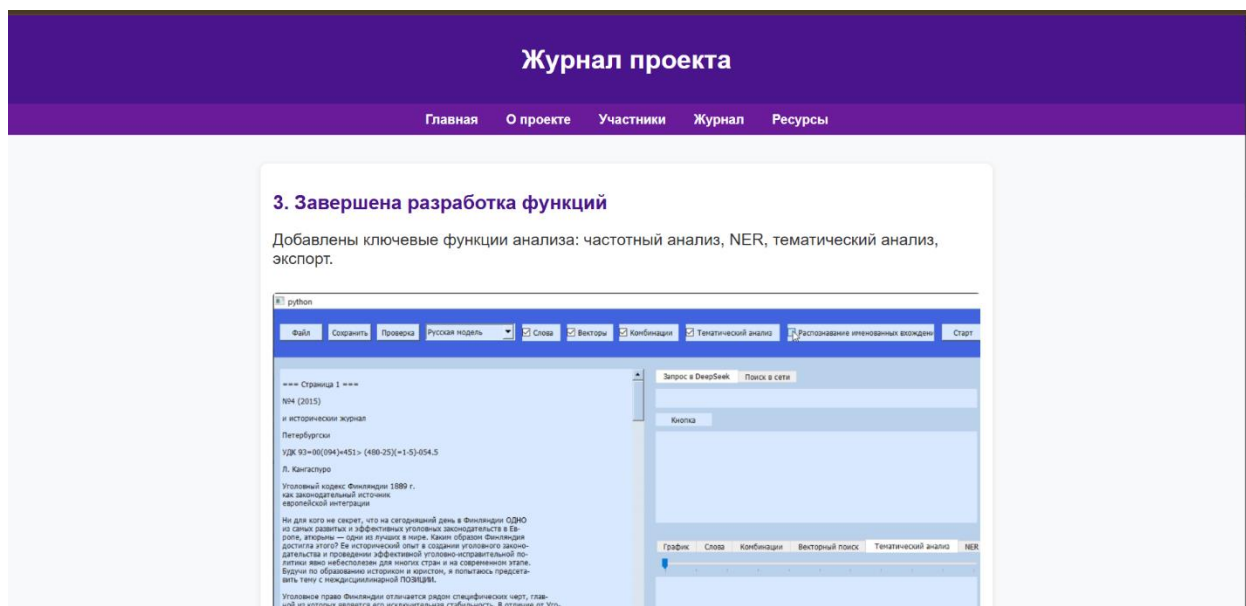
**Вклад в проект:**

- Ведение документации проекта
- Разработка интерфейса

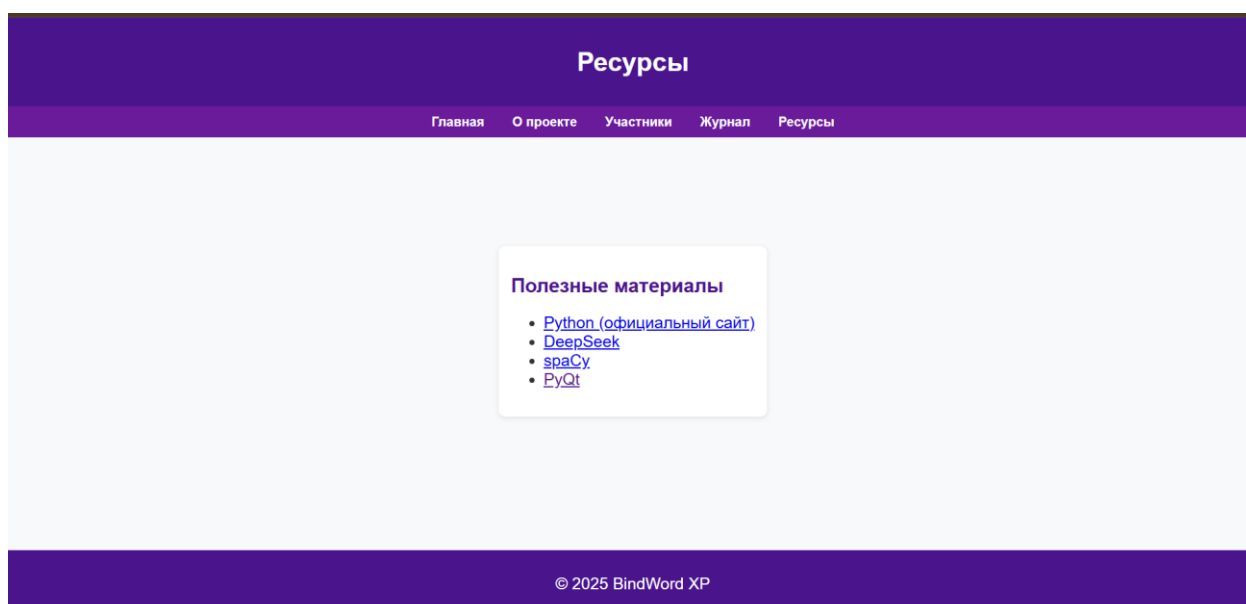
[Отчет по взаимодействию с партнером](#)

© 2025 BindWord XP

Страница «Журнал» с отражением процесса разработки программы



Страница «Ресурсы» отражает ссылки на используемые источники



3. Подготовлен отчет о взаимодействии с организацией-партнером в папке reports репозитория

## **ЗАКЛЮЧЕНИЕ**

В ходе прохождения проектной практики были закреплены ключевые навыки, необходимые для работы в области разработки и сопровождения программного обеспечения. Практическая деятельность включала освоение системы контроля версий Git, изучение языка разметки Markdown, а также разработку статического веб-сайта средствами HTML и CSS. Отдельное внимание уделялось вопросам взаимодействия с партнёрской организацией и участию в профильных мероприятиях, что позволило расширить профессиональные представления и связать полученные знания с реальными практиками индустрии.

Результатом выполненной работы стало формирование комплексного опыта, сочетающего теоретическую подготовку и практические умения. Освоенные инструменты и технологии создают основу для дальнейшего профессионального развития и применения в будущих учебных и исследовательских проектах, а также в практической деятельности в сфере информационных технологий.

## СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Информация о практике Московского политеха  
<https://mospolytech.ru/obuchauschimsya/praktika/?ysclid=m9fpo3pwmu710957340> (дата обращения: 15.09.2025).
2. Информация о проектной деятельности  
<https://mospolytech.ru/obuchauschimsya/proektnaya-deyatelnost/?ysclid=m9fpsda3ad786727228> (дата обращения: 15.09.2025).
3. Официальный сайт организации-партнера <https://mospolytech.ru/> (дата обращения: 15.09.2025).
4. Организационная структура организации-партнера  
<https://mospolytech.ru/sveden/struct/> (дата обращения: 15.09.2025).
5. Репозиторий GitHub, созданный в рамках проектной практики  
[https://github.com/lanshchekov/practice\\_2025\\_1](https://github.com/lanshchekov/practice_2025_1) (дата обращения: 15.09.2025).
6. Сайт проекта  
[https://lanshchekov.github.io/practice\\_2025\\_1/site/index.html](https://lanshchekov.github.io/practice_2025_1/site/index.html) (дата обращения: 15.09.2025).
7. Документация приложения BindWord XP  
[https://github.com/lanshchekov/practice\\_2025\\_1/blob/main/docs/bindwordxp\\_docs.pdf](https://github.com/lanshchekov/practice_2025_1/blob/main/docs/bindwordxp_docs.pdf) (дата обращения: 17.09.25)