**Документация, содержащая информацию, необходимую для эксплуатации экземпляра программы для ЭВМ**

**«Интеллектуальный модуль для классификации документов»**

Москва, 2025 г.

Содержание

[1 ОБЩИЕ ПОЛОЖЕНИЯ 3](#_Toc187940160)

[1.1 Общие сведения 3](#_Toc187940161)

[1.2 Информация, необходимая для установки и настройки 3](#_Toc187940162)

[1.3 Описание структуры 4](#_Toc187940163)

[2 НАЗНАЧЕНИЕ И УСЛОВИЯ ПРИМЕНЕНИЯ 4](#_Toc187940164)

[2.1 Назначение 4](#_Toc187940165)

[2.2 Условия применения 4](#_Toc187940166)

[**2.2.1 Аппаратные требования** 4](#_Toc187940167)

[**2.2.2 Программные требования** 4](#_Toc187940168)

[3 ПОДГОТОВКА К РАБОТЕ 5](#_Toc187940169)

[4 ОПИСАНИЕ РАБОТЫ 5](#_Toc187940170)

[4.1 Установка и настройка ПО 5](#_Toc187940171)

[4.2 Установка обновлений 5](#_Toc187940172)

[4.3 Штатное функционирование 6](#_Toc187940173)

[4.4 Резервное копирование и восстановление данных 6](#_Toc187940174)

[4.5 Проведение диагностики системы 6](#_Toc187940175)

[5 АВАРИЙНЫЕ СИТУАЦИИ 6](#_Toc187940176)

[6 ЭКСПЛУАТАЦИЯ 6](#_Toc187940177)

[6.1 Подготовка к работе 6](#_Toc187940178)

[6.2 Использование системы по назначению 7](#_Toc187940179)

[6.3 Завершение работы системы 11](#_Toc187940180)

# **1 ОБЩИЕ ПОЛОЖЕНИЯ**

## **Общие сведения**

Программа для ЭВМ «Интеллектуальный модуль для классификации документов» является средством для автоматизации классификации судебных документов с использованием машинного обучения. Основные функции: классификация судебных документов на основе их содержания; определение типа документа; автоматическое перемещение документа в соответствующую папку в зависимости от класса. Программа предназначена для использования в коллекторской системе и может быть интегрирована с существующими системами документооборота. Программа обеспечивает высокую точность классификации судебных документов (достигая accuracy 0.976 на тестовой выборке), что позволяет ускорить и упростить процесс классификации и регистрации судебных документов. Программа также может быть использована в других областях, где требуется классификация документов на основе их содержания.

## **Информация, необходимая для установки и настройки**

Действия по настройке и установке ПО программы не требуют специальных технических знаний и могут быть выполнены пользователем самостоятельно. Однако, для корректной работы программы необходимо выполнить следующие шаги:

1. Установить и добавить переменные среды для Tesseract и Poppler
2. Перезапустить систему для применения изменений

После выполнения этих шагов программа будет готова к работе.

## **1.3 Описание структуры**

Программа состоит из следующих компонентов:

* + Модуль обработки PDF-файлов
  + Модуль обученной модели классификации текстов
  + Модуль сортировка документов по категориям

Среда разработки ПО:

* Изолированная подсеть на основе ОС Windows.

Для корректной работы с платформой необходима следующая конфигурация автоматизированного рабочего места пользователя:

* Минимальные требования к системе - 4ядра
  + 4Gb RAMM доступной памяти
  + 100Gb SSD

# **2 НАЗНАЧЕНИЕ И УСЛОВИЯ ПРИМЕНЕНИЯ**

## **2.1 Назначение**

Программа для ЭВМ «Интеллектуальный модуль для классификации документов» представляет собой универсальное программное решение для автоматизации классификации судебных документов с использованием машинного обучения.

## **2.2 Условия применения**

### **2.2.1 Аппаратные требования**

Языки программирования, применявшиеся при разработке ПО:

* Python 3.11

Среда разработки ПО:

* Jupyter Notebook.

Для корректной работы с платформой необходима следующая конфигурация автоматизированного рабочего места пользователя:

* Минимальные требования к системе - 4ядра
  + 4Gb RAMM доступной памяти
  + 100Gb SSD

### **2.2.2 Программные требования**

Поддерживаемые ОС:

* + Windows 10

Необходимое ПО сторонних производителей:

* Tesseract-OCR
* Poppler-21.03.0

# **3 ПОДГОТОВКА К РАБОТЕ**

* Для начала работы необходимо установить необходимое ПО сторонних производителей
* Для работы с программой необходимо иметь доступ к PDF-документам

# **4 ОПИСАНИЕ РАБОТЫ**

## **4.1 Установка и настройка ПО**

Установка программы производится путем запуска готового файла exe. Для корректной работы программы необходимо выполнить следующие шаги:

1. **Добавить переменные среды для Tesseract и Poppler:**
   * Перейти в раздел "Свойства системы":
     + Нажать Win + Pause/Break или щелкнуть правой кнопкой мыши на "Компьютер" и выбрать "Свойства".
   * Выбрать "Дополнительные параметры системы".
   * Перейти в раздел "Переменные среды".
2. **Добавить новую переменную среды для Tesseract:**
   * Найти переменную Path в списке системных переменных и выбрать ее.
   * Нажать "Изменить", затем "Создать" или "Добавить" (в зависимости от версии Windows).
   * Ввести путь к папке с файлами Tesseract (например, C:\Program Files\Tesseract-OCR).
3. **Добавить новую переменную среды для Poppler:**
   * Повторить шаги для добавления пути к папке с файлами Poppler (например, C:\Program Files\poppler\bin).
4. **Перезапустить систему для применения изменений.**

После выполнения этих шагов программа будет готова к работе.

## **4.2 Установка обновлений**

Установка обновлений производится правообладателем. Обновление ПО пользователем самостоятельно невозможна, так как требует глубокого знания на уровне кода. В ряде случаев обновление может производиться удаленно, по договоренности с пользователем.

## **4.3 Штатное функционирование**

Программа функционирует в режиме обработки PDF-файлов, который включает в себя следующие этапы:

* Обработка PDF-файлов с помощью модуля обработки PDF-файлов
* Классификация текстов с помощью модуля программного модуля.

## **4.4 Резервное копирование и восстановление данных**

Резервное копирование данных производится путем создания копии файла конфигурации и данных, которые используются программой.

Восстановление данных производится путем восстановления файла конфигурации и данных из резервной копии.

## **4.5 Проведение диагностики системы**

При эксплуатации ПО может производиться удаленная или очная диагностика ПО правообладателем. Самостоятельная диагностика пользователем возможна только на уровне использования мониторингового ПО, в других случаях, невозможна, так как требует глубоких знаний на уровне кода.

# **5 АВАРИЙНЫЕ СИТУАЦИИ**

1. Информацию об аварийных ситуациях можно получить через лог-файлы программы
2. В случае ошибок в работе программы необходимо обратиться в техподдержку

# **6 ЭКСПЛУАТАЦИЯ**

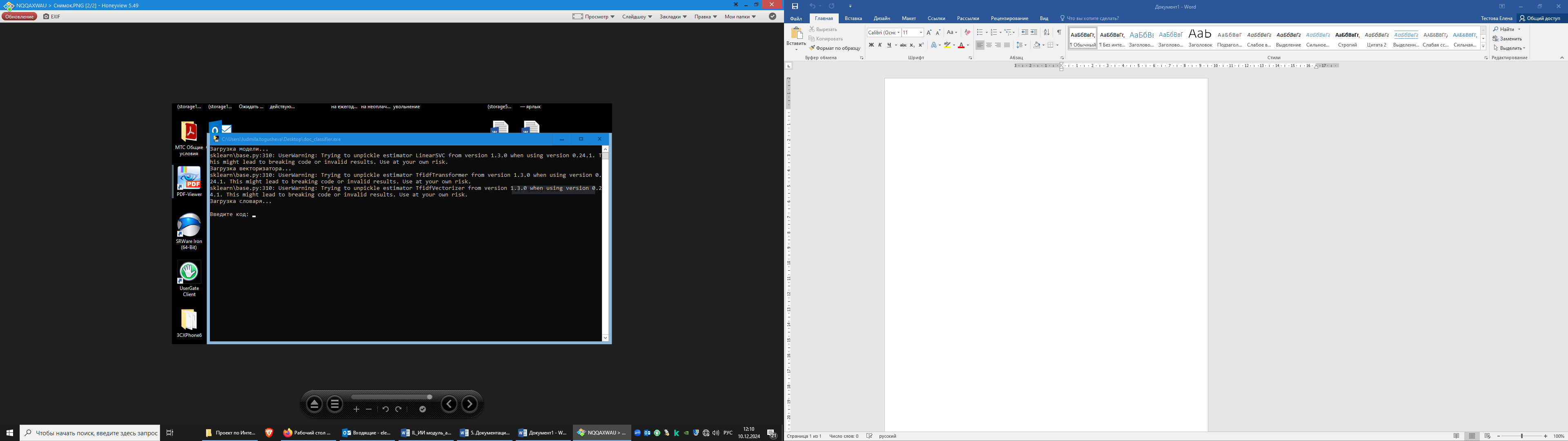
## **6.1 Подготовка к работе**

* Для начала работы необходимо установить необходимое ПО сторонних производителей
* Для работы с программой необходимо иметь доступ к PDF-документам

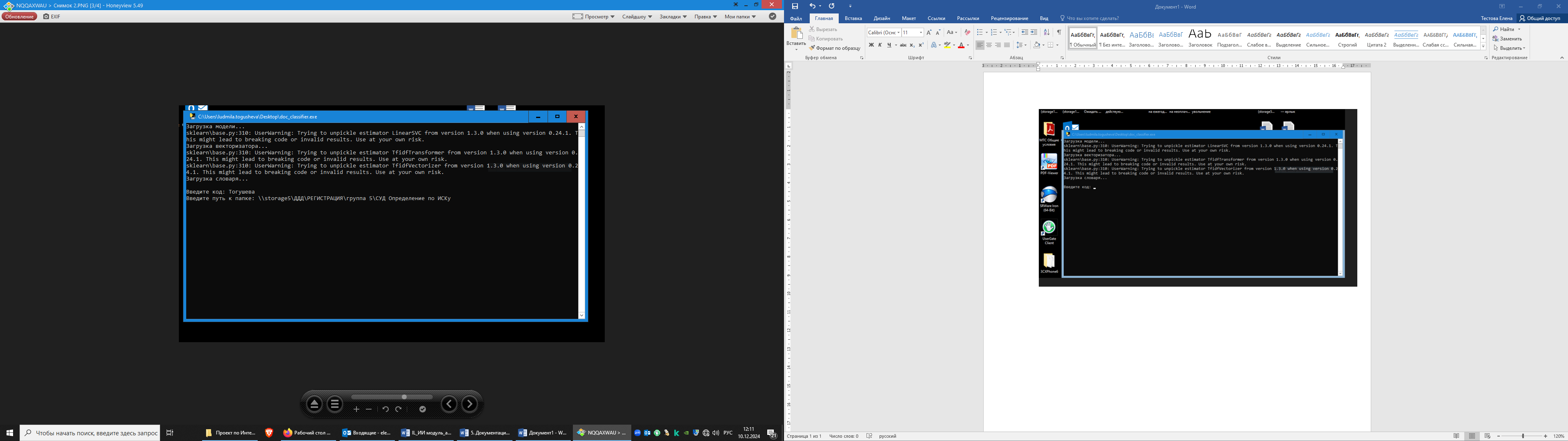
## **6.2 Использование системы по назначению**

Метод представляет из себя следующее:

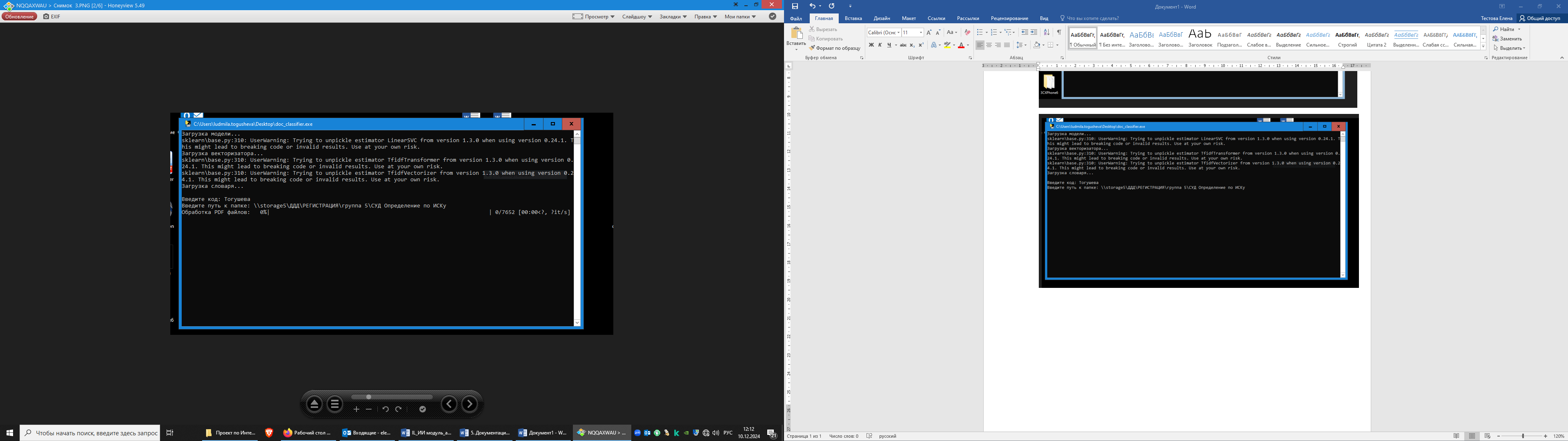
* Ввод кода сотрудника, обрабатывающего PDF документов

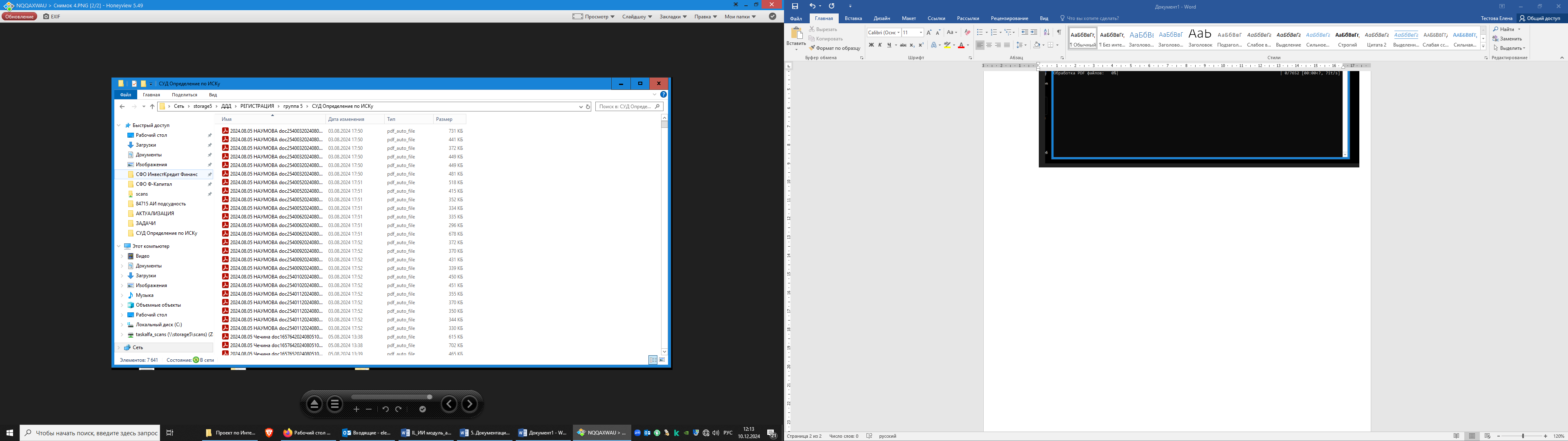


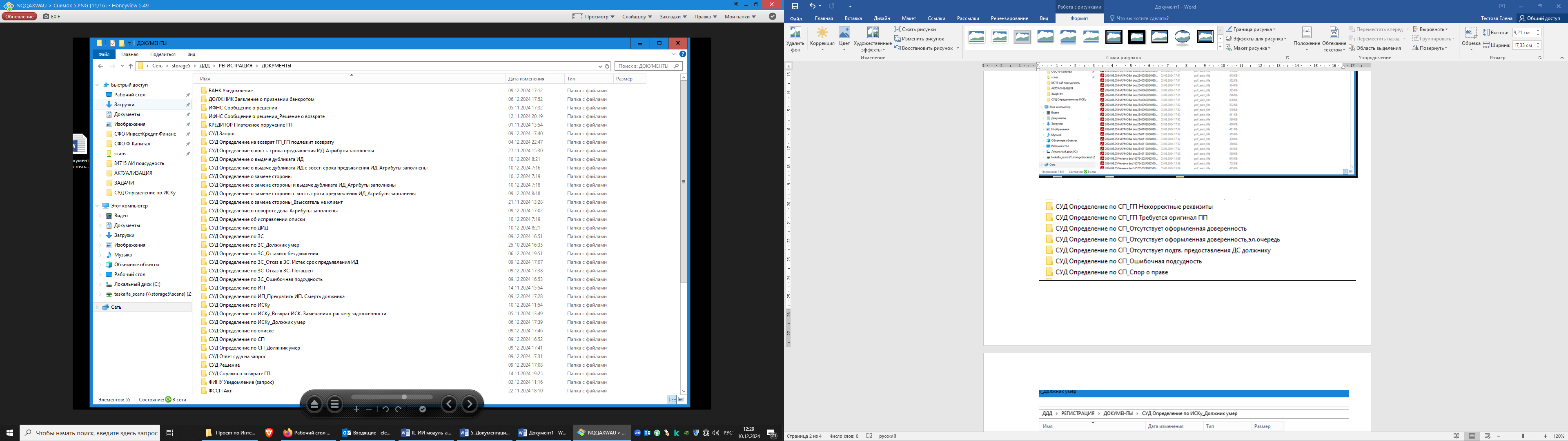
* Ввод ссылки папки, обрабатываемых PDF документов

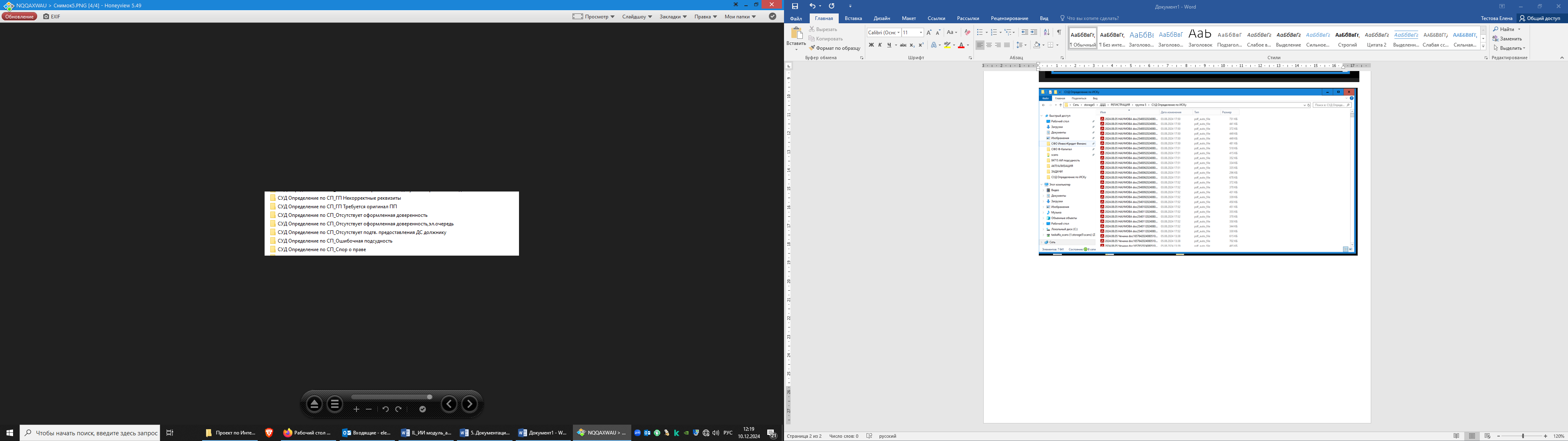


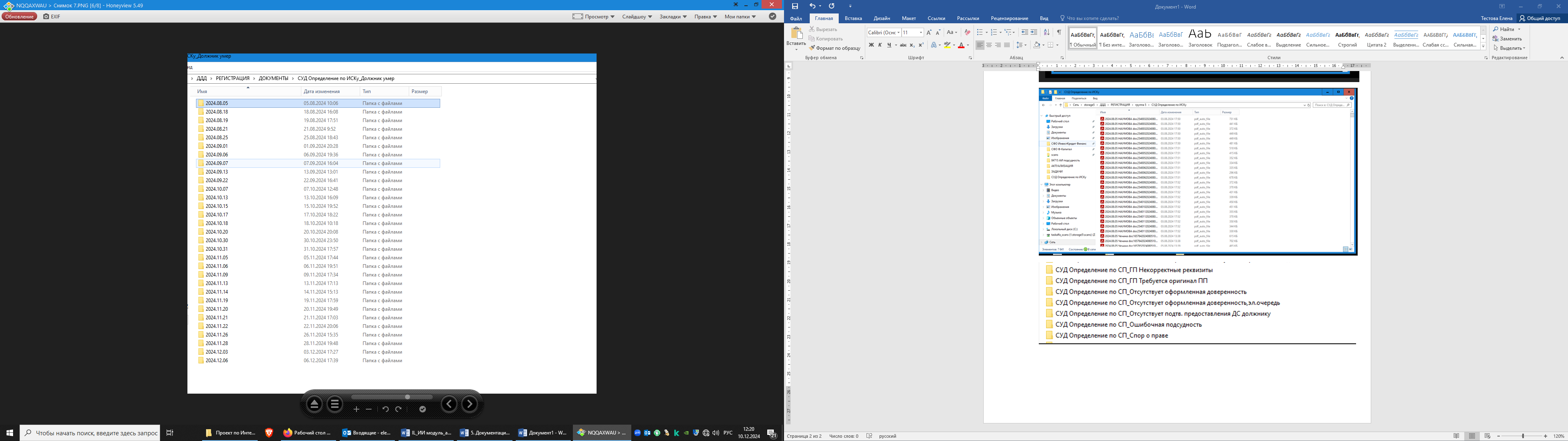
* Загрузка с последующей обработкой и сортировкой PDF-документов

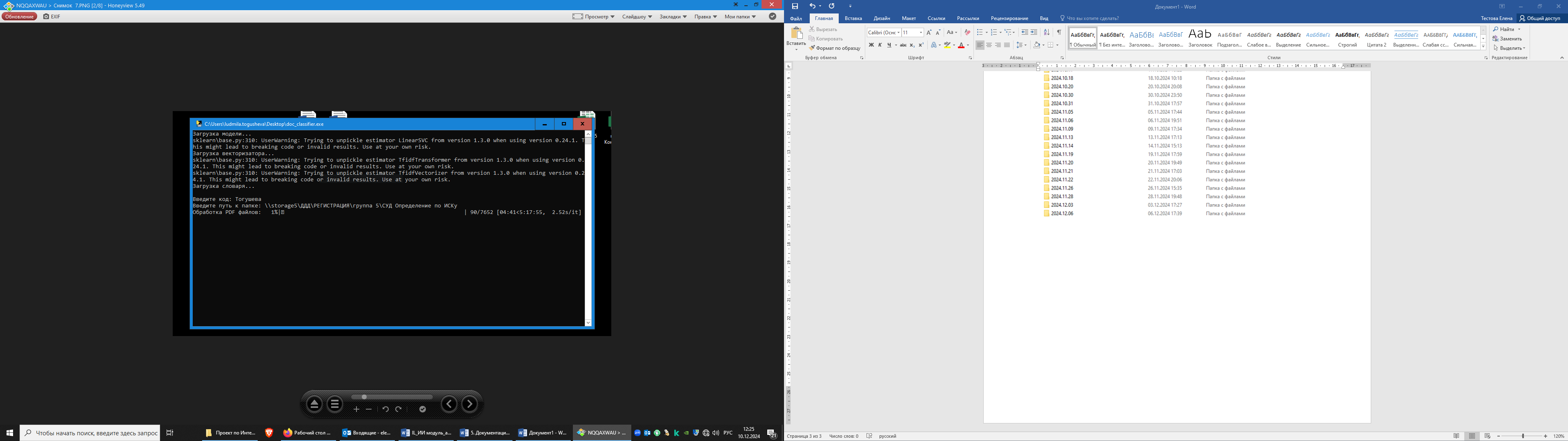












## **6.3 Завершение работы программы**

Программа завершает работу после обработки всех PDF документов:

