

**ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

Факультет компьютерных наук
Образовательная программа бакалавриата «Программная инженерия»

СОГЛАСОВАНО
Старший преподаватель департамента
программной инженерии

_____/А. Д. Игнатов /
«__» _____ 2026 г.

УТВЕРЖДЕНО
Академический руководитель
образовательной программы
«Программная инженерия»,
старший преподаватель департамента
программной инженерии
_____/ Н. А. Павлочев /
«__» _____ 2024 г.

Подп. и дата	
Инв. № дубл.	
Взам. инв. №	
Подп. и дата	
Инв. № подл	

**Система распознавания PDF и изображений
для преобразования данных в структурированный формат**

Текст программы
ЛИСТ УТВЕРЖДЕНИЯ
RU.17701729.05-13 ТЗ 01-1

Исполнитель:
студент группы БПИ243
_____/ И. А. Ковалев /
«__» _____ 2026 г.

Исполнитель:
студент группы БПИ243
_____/ А. А. Торосян /
«__» _____ 2026 г.

**Система распознавания PDF и изображений
для преобразования данных в структурированный формат**

Инев. № подл	Подп. и дата	Взам. инв. №	Инев. № дубл.	Подп. и дата

Текст программы
RU.17701729.05-13 ТЗ 01.1
Листов 7

АННОТАЦИЯ

Настоящий документ «Текст программы» для проекта «Система распознавания PDF и изображений для преобразования данных в структурированный формат» содержит следующие разделы: «Введение», «Текст программы». Программа написана на языке программирования Python версии 3.9+ с применением Flask. Среда разработки – PyCharm.

В разделе «Введение» приведены краткие и полные названия продукта.

В разделе «Текст программы» содержатся описания классов, а также ссылка на репозиторий с кодом.

Настоящий документ разработан в соответствии с требованиями:

- 1) ГОСТ 19.101-77 Виды программ и программных документов;
- 2) ГОСТ 19.106-78 Требования к программным документам, выполненным печатным способом;
- 3) ГОСТ 19.103-77 Обозначения программ и программных документов;
- 4) ГОСТ 19.104-78 Основные надписи;
- 5) ГОСТ 19.105-78 Общие требования к программным документам;
- 6) ГОСТ 19.602-78 Правила дублирования, учета и хранения программных документов, выполненных печатным способом;

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.05-13 ТЗ 01.1				
Инв. № подл.	Подп. и дата	Взам. Инв. №	Инв. № дубл.	Подп. и дата

СОДЕРЖАНИЕ

Оглавление

1. ВВЕДЕНИЕ.....	4
1.1. Наименование	4
1.2. Краткая характеристика области применения	4
2. ТЕКСТ ПРОГРАММЫ	5
ПРИЛОЖЕНИЕ.....	6
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	6
ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ.....	7

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.05-13 ТЗ 01.1				
Инв. № подл.	Подп. и дата	Взам. Инв. №	Инв. № дубл.	Подп. и дата

1. ВВЕДЕНИЕ

1.1. Наименование

Наименование программы – «Система распознавания PDF и изображений для преобразования данных в структурированный формат»

Наименование программы на английском языке – «PDF And Image Recognition System For Structured Data Conversion»

1.2. Краткая характеристика области применения

Данное приложение предназначено для автоматизации процесса распознавания данных в PDF-документах и изображениях с последующим преобразованием их в структурированные форматы. Система позволяет пользователям загружать документы в различных форматах, автоматически извлекать из них данные, анализировать структуру содержимого и преобразовывать результаты в удобные для дальнейшей обработки форматы.

Основная цель системы – упростить и автоматизировать процесс обработки неструктурированных данных, содержащихся в PDF-документах и изображениях, путем их преобразования в структурированный вид для последующего анализа, хранения или интеграции с другими информационными системами.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.05-13 ТЗ 01.1				
Инв. № подл.	Подп. и дата	Взам. Инв. №	Инв. № дубл.	Подп. и дата

2. ТЕКСТ ПРОГРАММЫ

Репозиторий программы находится по ссылке: <https://github.com/Restmine20/PDF-And-Image-Recognition-System-For-Structured-Data-Conversion>.

Основные модели данных находятся в папке src/models.

Коды для клиентской части находятся в папке src/frontend.

Коды для серверной части находятся в папке src/backend.

Полный текст не приводится в связи с большим количеством строк кода.

Изм.	Лист	№ докум.	Подп.	Дата
RU.17701729.05-13 ТЗ 01.1				
Инв. № подл.	Подп. и дата	Взам. Инв. №	Инв. № дубл.	Подп. и дата

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1) ГОСТ 19.101–77 Виды программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
- 2) ГОСТ 19.102–77 Стадии разработки. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
- 3) ГОСТ 19.103–77 Обозначения программ и программных документов. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
- 4) ГОСТ 19.104–78 Основные надписи. // Единая система программной документации. – М.: ИПК Издательство стандартов, 2001;
- 5) ГОСТ 19.105–78 Общие требования к программным документам. // Единая система программной документации. – М.: ИПК Издательство стандартов, 2001;
- 6) ГОСТ 19.106–78 Требования к программным документам, выполненным печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
- 7) ГОСТ 19.602–78 ЕСПД. Правила дублирования, учета и хранения программных документов, выполненных печатным способом. — Москва: Стандартинформ, 2005
- 8) ГОСТ 19.603–78 Общие правила внесения изменений. // Единая система программной документации. – М.: ИПК Издательство стандартов, 2001;

ЛИСТ РЕГИСТРАЦИИ ИЗМЕНЕНИЙ

[illegible]