

**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. Павлов /  
«\_\_\_» \_\_\_\_\_ 2024 г.

<i>Инв. № подп</i>	<i>Подп. и дата</i>	<i>Инв. № дубл.</i>	<i>Подп. и дата</i>

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

**Текст программы**

**ЛИСТ УТВЕРЖДЕНИЯ**

**RU.17701729.05-13 Т3 01-1**

Исполнитель:

студент группы БПИ243  
\_\_\_\_\_ / И. А. Ковалев /  
«\_\_\_» \_\_\_\_\_ 2026 г.

Исполнитель:

студент группы БПИ243  
\_\_\_\_\_ / А. А. Торосян /  
«\_\_\_» \_\_\_\_\_ 2026 г.

**2026**

УТВЕРЖДЕН

RU.17701729.05-13 Т3 01.1

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

<i>Инв. № подп</i>	<i>Подп. и дата</i>	<i>Взам. инв. №</i>	<i>Инв. № дубл.</i>	<i>Подп. и дата</i>

**Текст программы**

**RU.17701729.05-13 Т3 01.1**

**Листов 7**

**2026**

## 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**

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

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 Т3 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 Т3 01.1				
Инв. № подл.	Подп. и дата	Взам. Инв. №	Инв. № дубл.	Подп. и дата

## **ПРИЛОЖЕНИЕ**

### **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