|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

ТЕХНИЧЕСКОЕ ЗАДАНИЕ

на разработку

программного обеспечения для анализа и обработки речевых сигналов

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

2025

Содержание

[Введение 2](#_Toc210254043)

[1 Термины и определения 3](#_Toc210254044)

[2 Перечень сокращений 4](#_Toc210254045)

[3 Основные сведения о разработке 5](#_Toc210254046)

[3.1 Наименование разработки 5](#_Toc210254047)

[3.2 Цель и задачи 5](#_Toc210254048)

[3.3 Сведения об участниках разработки 5](#_Toc210254049)

[3.4 Сроки разработки 5](#_Toc210254050)

[3.5 Назначение разработки 6](#_Toc210254051)

[3.5.1 Функциональное назначение 6](#_Toc210254052)

[3.5.2 Эксплуатационное назначение 6](#_Toc210254053)

[4 Описание предметной области 7](#_Toc210254054)

[5 Требования к результатам разработки 11](#_Toc210254059)

[5.1 Требования к функциональным характеристикам 11](#_Toc210254060)

[5.2 Требования к показателям назначения 11](#_Toc210254061)

[5.3 Требования к технологическому стеку 11](#_Toc210254062)

[5.4 Требования к пользовательскому интерфейсу 12](#_Toc210254063)

[5.5 Требования к видам обеспечения 13](#_Toc210254064)

[5.5.1 Требования к математическому обеспечению 13](#_Toc210254065)

[5.5.2 Требования к информационному обеспечению 14](#_Toc210254066)

[5.5.3 Требования к метрологическому обеспечению 14](#_Toc210254067)

[5.5.4 Требования к техническому обеспечению 14](#_Toc210254068)

[5.6 Требования к надежности 14](#_Toc210254069)

[5.7 Требования к безопасности 15](#_Toc210254070)

[5.8 Требования к патентной чистоте 15](#_Toc210254071)

[5.9 Требования к перспективам развития 15](#_Toc210254072)

[6 Состав и содержание работ 16](#_Toc210254073)

[7 Требования к документированию 18](#_Toc210254074)

[8 Требования к приемно-сдаточным процедурам 19](#_Toc210254075)

# Введение

В данном документе представлено техническое задание на разработку программного обеспечения для анализа и обработки речевых сигналов.

Настоящий документ содержит подробное описание всех аспектов проекта, включая цели и задачи разработки, требования к возможностям и интерфейсу, а также план работ.

Документ предназначен для исполнителя, который будет работать над проектом, а также для заказчика, чтобы контролировать ход разработки и убедиться, что результат работ соответствует содержанию документа.

# Термины и определения

В данном разделе представлен список терминов и определений, содержащихся в настоящем документе.

Речевой сигнал – акустический сигнал, порождаемый речевым аппаратом человека и используемый для передачи информации в устной форме.

Фонема – минимальная единица звуковой системы языка, которая различает значения слов.

Форманта – акустическая характеристика звуков речи (прежде всего гласных), связанная с уровнем частоты голосового тона и образующая тембр звука.

Транскрибирование – запись речи в виде последовательности фонем или звуков.

Сегментация – разбиение звукового сигнала на части (слоги, фонемы, слова).

# Перечень сокращений

В данном разделе представлен перечень сокращений, используемых в настоящем документе.

ПО – программное обеспечение.

ПК – персональный компьютер.

БД – база данных.

ОЗУ – оперативное запоминающее устройство.

ГБ – гигабайт.

ГГц – гигагерц.

ФЗ – федеральный закон.

ГОСТ – государственный стандарт.

ПЭВМ – персональная электронно-вычислительная машина.

ТЗ – техническое задание.

# Основные сведения о разработке

## Наименование разработки

Наименование данной разработки – программное обеспечение для анализа и обработки речевых сигналов.

## Цель и задачи

Целью разработки является создание программного средства для автоматизации анализа речевых сигналов, позволяющего сократить трудоёмкость и повысить точность при исследовании акустических характеристик речи.

Результатом внедрения системы станет снижение временных затрат специалистов на ручной анализ аудиозаписей и повышение достоверности получаемых данных. Для достижения цели необходимо:

* обеспечить возможность импорта и воспроизведения аудиозаписей в стандартных форматах (WAV, MP3, AAC, FLAC);
* реализовать базовые инструменты предварительной обработки звуковых сигналов (нормализация громкости, фильтрация шумов, обрезка и выделение фрагментов);
* разработать алгоритмы автоматической сегментации аудио на фонемы;
* реализовать функции классификации фонем (гласные, согласные) и определения ударных гласных;
* создать модуль поиска звуковых сочетаний с отображением результатов на спектрограмме и формированием списка с временными метками;
* предусмотреть экспорт результатов анализа в распространённых форматах для дальнейшего использования.

## Сведения об участниках разработки

Исполнителем является студент Колледжа ФГБОУ ВО «Вятский государственный университет» учебной группы ИСПк-403-52-00 Бочкарёва Виктория Дмитриевна.

Заказчиком является Колледж ФГБОУ ВО «Вятский государственный университет», представленный коллективом преподавателей в составе:

* Долженкова Мария Львовна – преподаватель по учебной практике УП.03.

## Сроки разработки

Начало разработки – 18.09.2025.

Конец разработки – ХХ.11.2025.

## Назначение разработки

В данном разделе описаны функциональное и эксплуатационное назначение разрабатываемой системы.

### Функциональное назначение

Функциональное назначение ПО заключается в автоматизации процессов анализа и обработки речевых сигналов: выделение речевых сегментов, идентификация фонетических структур, проведение спектрального анализа, а также формирование отчетов для дальнейшего исследования и сравнения результатов.

### Эксплуатационное назначение

Система может быть использована исследователями, преподавателями и студентами в области лингвистики, акустики и цифровой обработки сигналов для изучения характеристик речи, а также для проведения практических и научных работ, связанных с анализом речевых данных.

# Описание предметной области

Обработка и анализ речевых сигналов играют фундаментальную роль в таких областях, как лингвистика, криминалистика, медицина и разработка систем распознавания речи. Однако эффективная работа с аудиоданными требует значительных усилий и специальных знаний. В последние десятилетия процесс анализа речевой информации претерпел серьезные изменения благодаря развитию технологий цифровой обработки сигналов.

Традиционные подходы к анализу аудиозаписей, такие как ручное прослушивание, транскрибирование и сегментация, являются чрезвычайно трудоемкими процессами. Ручной поиск определенных звуковых сочетаний в длинных записях мог занимать многие часы, а такие задачи, как фильтрация шумов или нормализация громкости, требовали использования сложного студийного оборудования. Все это сопровождалось высокой вероятностью ошибок, связанных с человеческим фактором, и субъективностью в интерпретации данных.

С появлением цифровых аудиоредакторов в конце XX века стало возможным частичное облегчение этих процессов. Однако лишь с развитием специализированных алгоритмов и увеличением вычислительных мощностей началась активная фаза внедрения программных решений для глубокого фонетического анализа, что позволило радикально изменить структуру работы исследователей.

В наше время автоматизация стала ключевым элементом в работе с речевыми данными, что привело к значительному сокращению временных затрат и повысило общую эффективность исследований. Именно здесь на передний план выходит роль программного обеспечения для обработки и анализа речевых сигналов.

Программное обеспечение для обработки и анализа речевых сигналов – это специализированная программа, предназначенная для автоматизации и оптимизации процессов работы с аудиозаписями речи.

Внедрение данного программного обеспечения преследует следующие цели:

* ускорение таких операций, как импорт файлов, нормализация громкости и фильтрация шумов;
* автоматическая сегментация речи на фонемы, определение их типов (гласные, согласные) и выделение ударных гласных;
* предоставление инструментов для быстрого поиска звуковых сочетаний и их визуализации на спектрограмме;
* автоматическое создание списков найденных сочетаний с точными временными метками для дальнейшего использования;
* упрощение процессов обрезки, выделения нужных фрагментов и экспорта данных для научных публикаций или практического применения;
* сокращение временных затрат на рутинные технические задачи и концентрация на интерпретации результатов анализа.

## Обзор аналогов

На рынке существует ряд, которые могут быть использованы в качестве аналогов.

## Аналог №1 – «Praat»

Praat – это бесплатная программа для анализа и синтеза речи, разработанная в 1991 году Полом Боэрсма и Давидом Венинком. Она широко применяется в фонетике и лингвистике, предоставляет возможности спектрального анализа, ручной разметки и работы с формантами.

Интерфейс программы представлен на рисунке 4.1.

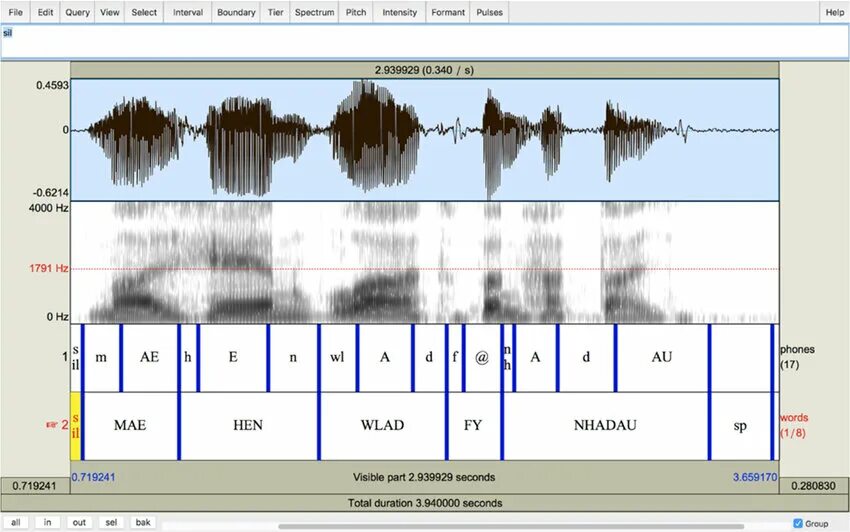


Рисунок 4.1 – Интерфейс программы «Praat»

Преимущества:

* программа бесплатна и активно используется в научных исследованиях;
* предоставляет большой набор инструментов для спектрального анализа;
* поддерживает русский язык в обработке аудиофайлов.

Недостатки:

* интерфейс устаревший и сложный для начинающих пользователей;
* отсутствует автоматическое определение фонем, гласных и согласных;
* сегментация речи выполняется вручную, что требует больших временных затрат.

## Аналог №2 – «Phonexia Speech Technologies»

Phonexia Speech Technologies – это коммерческий продукт, разрабатываемый одноимённой компанией с 2006 года. Система используется для распознавания речи, биометрической идентификации и анализа аудио. Имеет модуль Phoneme Recogniser, выполняющий разметку звуковых единиц.

Интерфейс программы представлен на рисунке 4.2.

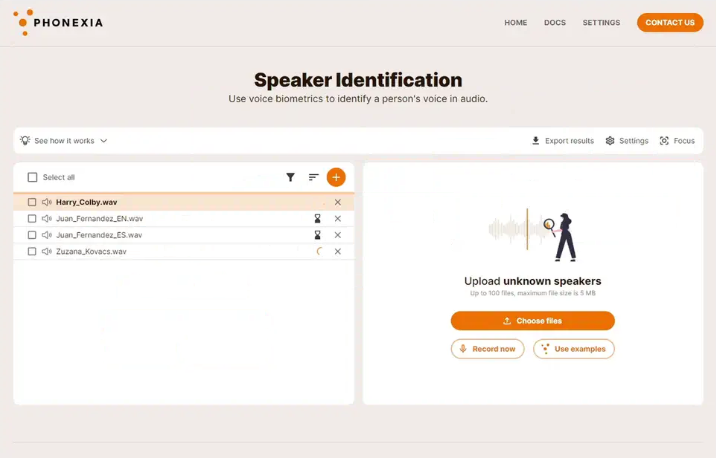


Рисунок 4.2 – Интерфейс программы «Phonexia Speech Technologies»

Преимущества:

* современный интерфейс и поддержка интеграции с другими системами;
* возможность автоматической работы с аудио и выделения фонем;
* применяется в криминалистике и корпоративной аналитике.

Недостатки:

* ограниченная поддержка русского языка;
* продукт является коммерческим и требует лицензии;
* программа не выполняет автоматического различения гласных, согласных и ударных фонем.

## Аналог №3 – «Gruut»

Gruut – это открытое программное обеспечение для фонематической транскрипции, разработанное сообществом энтузиастов. Оно предназначено в первую очередь для преобразования текста в последовательности фонем, что используется в системах синтеза речи.

Преимущества:

* полностью бесплатная и открытая лицензия;
* поддерживает множество языков, включая русский;
* интегрируется с проектами синтеза речи (например, TTS).

Недостатки:

* работает только по схеме «текст → фонемы», не анализирует аудио;
* отсутствует графический интерфейс, работа ведётся через код.

# Требования к результатам разработки

В этом разделе описываются требования к разработке. Программное обеспечение должно соответствовать представленным в данном разделе критериям.

## Требования к функциональным характеристикам

Программа должна осуществлять выполнение следующих функций:

* импорт файлов в стандартных форматах (WAV, AAC, MP3, FLAC);
* воспроизведение аудиозаписей внутри приложения;
* нормализация уровня громкости;
* фильтрация шумов;
* обрезка и выделение нужных фрагментов;
* сегментация аудио на фонемы;
* определение гласных и согласных фонем;
* определение ударных гласных;
* поиск звуковых сочетаний;
* отображение найденных сочетаний на спектрограмме;
* формирование списка сочетаний с временными метками;
* экспорт результатов.

## Требования к показателям назначения

Информационная система должна обеспечивать следующие показатели назначения:

* время на полный запуск системы не должно превышать 30 секунд;
* время отклика при выполнении основных операций (загрузка аудиофайла, построение спектрограммы, расчет характеристик) не должно превышать 3 секунд для файлов длительностью до 10 секунд;
* максимальная поддерживаемая длительность анализируемого сигнала — до 60 минут.

## Требования к технологическому стеку

При разработке программы должен использоваться язык программирования **C++**.

Иные требования к технологическому стеку не предъявляются.

## Требования к пользовательскому интерфейсу

Пользовательский интерфейс должен быть интуитивно понятен и обеспечивать все необходимые возможности программы.

При запуске программы открывается стартовый экран с возможностью создать новый проект, загрузив аудио, или вернуться к старым проектам. Прототип окна стартового экрана представлен на рисунке 5.4.1.

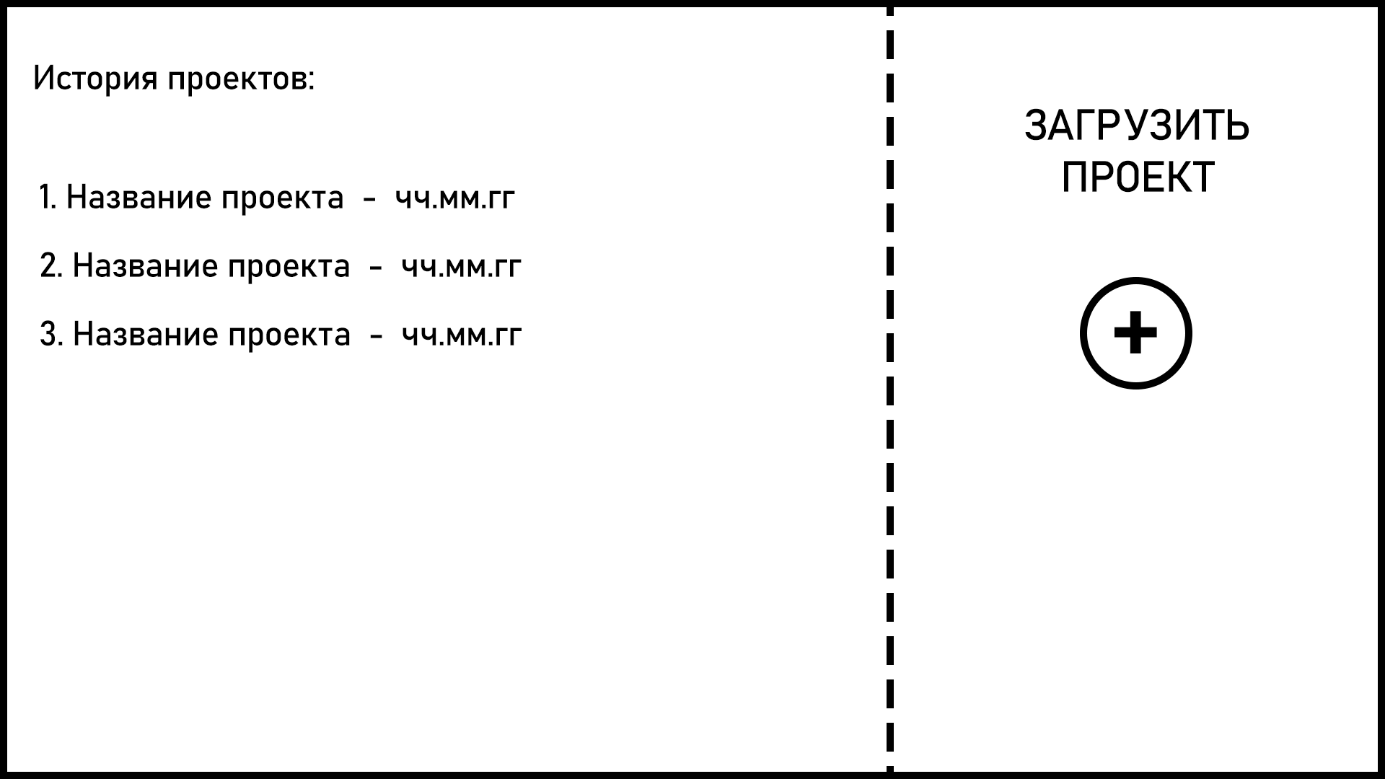


Рисунок 5.4.1 – Прототип экранной формы стартового окна

На стартовом экране предоставляются возможности:

* просматривать список ранее созданных проектов с указанием названия и даты создания/последнего изменения;
* загружать существующий проект из списка;
* создавать новый проект с помощью кнопки «Загрузить проект» (с иконкой «+»).

Для работы с выбранным проектом и аудиофайлами программа должна предоставлять главный экран, являющийся основным рабочим пространством пользователя. Прототип окна главного экрана представлен на рисунке 5.4.2.

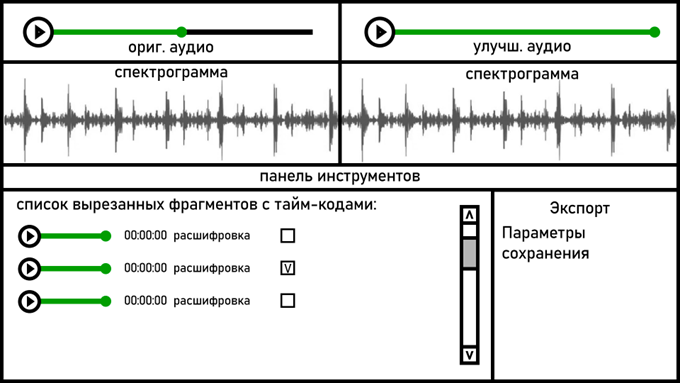


Рисунок 5.4.2 – Прототип экранной формы главного окна

В рабочем пространстве представлены следующие возможности:

* воспроизводить оригинальное аудио и улучшенное аудио через встроенные плееры;
* просматривать спектрограмму для оригинального и обработанного аудио;
* управлять воспроизведением (старт, пауза, перемотка, регулировка громкости);
* вырезать и сохранять отдельные аудиофрагменты с тайм-кодами;
* просматривать список найденных фрагментов аудио;
* прослушать каждый фрагмент отдельно;
* отмечать найденные фрагменты для экспорта;
* экспортировать результаты в выбранном формате.

## Требования к видам обеспечения

В данном подразделе содержатся требования к видам обеспечения разработки.

### Требования к математическому обеспечению

Для реализации заявленных функций должны быть использованы следующие математические модели и алгоритмы цифровой обработки сигналов: нормализация уровня громкости, фильтрация шумов, сегментация и анализ фонем, поиск звуковых сочетаний.

### Требования к информационному обеспечению

В данном пункте содержатся требования к информационному обеспечению разработки.

#### Требования к форматам хранения данных

Система должна обеспечивать импорт аудиофайлов в стандартных форматах несжатого аудио (WAV) и сжатого аудио с потерями и без потерь (MP3, AAC, FLAC).

Результаты анализа, включая фонетическую транскрипцию, список найденных сочетаний с временными метками и параметры сегментации, должны храниться в структурированных текстовых форматах.

#### Требования к лингвистическому обеспечению

Информационная система должна обеспечивать полную поддержку русского языка, включая локализацию интерфейса, корректную обработку аудио данных на русском языке, а также предоставление документации на русском языке.

### Требования к метрологическому обеспечению

Требования к метрологическому обеспечению не предъявляются.

### Требования к техническому обеспечению

Минимальные технические требования к запуску ИС включают следующие компоненты персонального компьютера:

* процессор с частотой 2Ггц;
* ОЗУ 8 ГБ;
* монитор;
* мышь или тачпад;
* клавиатура;
* операционная система Windows 11.

## Требования к надежности

Надежное функционирование ИС должно быть обеспечено выполнением организационно-технических мероприятий, указанных ниже:

* организацией бесперебойного питания технических средств;
* осуществлением контроля входных данных;
* регулярным следованием рекомендациям Министерства труда и социального развития РФ, изложенных в Постановлении от 23 июля 1998 г. «Об утверждении  
  межотраслевых типовых норм времени на работы по сервисному обслуживанию  
  ПЭВМ и оргтехники и сопровождению программных средств».

## Требования к безопасности

Безопасное функционирование программы должно обеспечиваться комплексом организационно-технических мер, соответствующих ГОСТ 51188–98 «Защита информации. Испытания программных средств на наличие компьютерных вирусов».

Программа может работать с аудиозаписями, которые могут содержать персональные данные пользователей (голосовые сообщения, имена файлов, метаданные).

Сбор и использование таких данных допускается исключительно для целей анализа и обработки речевых сигналов.

Программа должна обеспечивать конфиденциальность информации, исключая её передачу третьим лицам без согласия пользователя.

## Требования к патентной чистоте

Программа не должна использовать результаты интеллектуального труда сторонних субъектов и нарушать права третьих лиц в соответствии с частью четвёртой Гражданского кодекса Российской Федерации от 18.12.2006 № 230-ФЗ.

В случае использования внешних библиотек или сервисов необходимо:

* применять их исключительно в рамках действующих лицензий или пользовательских соглашений;
* указывать авторство и условия использования в сопровождающей документации;
* обеспечивать, чтобы использование данных решений не нарушало патентную чистоту разрабатываемого программного обеспечения.

## Требования к перспективам развития

Программа может иметь следующие направления дальнейшего развития:

* расширение функционала за счёт внедрения методов машинного обучения и нейронных сетей для автоматической классификации речевых сигналов;
* интеграция с системами автоматического распознавания речи;
* разработка мобильной версии приложения для Android и iOS;
* расширение перечня поддерживаемых форматов аудио и отчётности.

# Состав и содержание работ

В рамках разработки программного продукта в соответствии с настоящим документом необходимо выполнить перечень работ, представленных в таблице 1.

Таблица 1 – Состав и содержание работ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № этапа | Наименование этапа | Длительность | Состав работ | Результат |
| 1 | Изучение предметной области и аналогов | 1 неделя | Анализ предметной области, изучение существующих программ для анализа и обработки речи, выявление их достоинств и недостатков | Сформировано представление о функционале будущей программы |
| 2 | Составление Технического задания | 1 неделя | Разработка и утверждение ТЗ, определение требований к функционалу и интерфейсу | Утверждённое техническое задание |
| 3 | Проектирование | 2 недели | Разработка архитектуры приложения, описание алгоритмов анализа речи, подготовка макетов интерфейса | Спроектированная структура ПО и макеты интерфейса |
| 4 | Реализация программного обеспечения | 5 недель | Написание исходного кода программы, реализация функций импорта аудио, спектрального анализа и поиска звуковых сочетаний | Рабочий прототип программы |
| 5 | Тестирование программы | 1 неделя | Проведение функционального тестирования, выявление ошибок и недоработок | Список ошибок и предложений по доработке |
| 6 | Доработка программы | 2 недели | Исправление выявленных ошибок, оптимизация работы интерфейса и алгоритмов | Исправленная версия программы |
| 7 | Составление отчетной документации | 1 неделя | В ходе работ происходит разработка и утверждение отчетной документации | Готовая отчётная документация |
| 8 | Приемно-сдаточные испытания | 1 неделя | Представление программы и документации комиссии, внесение результатов испытаний в ведомость | Итоговая оценка |

# Требования к документированию

Состав программной документации включает:

* техническое задание;
* руководство пользователя.

# Требования к приемно-сдаточным процедурам

Испытания проводятся комиссией, в лице заказчика, которым является Колледж ФГБОУ ВО «Вятский государственный университет», представленный преподавателем учебной практики УП.03 – Долженковой Марией Львовной.

Во время испытаний комиссия проверяет работу программы в соответствии со следующими позициями:

* набор функциональных тестов;
* корректное функционирование заданных в техническом задании функций;
* возможность функционирования на ПК с указанными минимальными системными требованиями.

Комиссии должны быть представлены эксплуатационные документы, разработанная программа и доклад.