# ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

# НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

# «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

**Факультет компьютерных наук**

**Департамент программной инженерии**

|  |  |
| --- | --- |
| СОГЛАСОВАНО  Доцент департамента больших данных и информационного поиска, кандидат компьютерных наук  М.Л. Каледин  «\_\_\_» \_\_\_\_\_\_\_\_\_\_2025 г. | УТВЕРЖДАЮ  Академический руководитель образовательной программы  «Программная инженерия» профессор департамента программной  инженерии, канд. техн. наук  Н.А. Павлочев  « » 2025 г. |

|  |  |
| --- | --- |
| *Подп. и дата* |  |
| *Инв. № дубл.* |  |
| *Взам. инв. №* |  |
| *Подп. и дата* |  |
| *Инв. № подл* | RU.17701729.12.10-01 ТЗ 01-1-ЛУ |

x

**Приложение для разделения аудио на основе нейронных сетей**

**Техническое задание**

**ЛИСТ УТВЕРЖДЕНИЯ**

**RU.17701729.** **12.10-01 ТЗ 01-1-ЛУ**

Исполнитель

студент группы БПИ223

/ Иванов Г.Я

«\_\_» 2025 г.

УТВЕРЖДЕНО

RU.17701729.12.10-01 ТЗ 01-1-ЛУ

|  |  |
| --- | --- |
| *Подп. и дата* |  |
| *Инв. № дубл.* |  |
| *Взам. инв. №* |  |
| *Подп. и дата* |  |
| *Инв. № подл* | RU.17701729.12.10-01 ТЗ 01-1-ЛУ |

x

**Приложение для разделения аудио на основе нейронных сетей**

**Техническое задание**

**RU.17701729. 12.10-01 ТЗ 01-1-ЛУ**

**Листов 22**

**Содержание**

[АННОТАЦИЯ 3](#_Toc197998575)

[ГЛОССАРИЙ 4](#_Toc197998576)

[1. ВВЕДЕНИЕ 5](#_Toc197998577)

[1.1 Наименование программы 5](#_Toc197998578)

[1.2 Краткая характеристика области примения программы: 5](#_Toc197998579)

[2. ОСНОВАНИЯ ДЛЯ РАЗРАБОТКИ 6](#_Toc197998580)

[2.1 Документы, на основании которых ведётся разработка 6](#_Toc197998581)

[2.2 Наименование темы разработки 6](#_Toc197998582)

[3. НАЗНАЧЕНИЕ РАЗРАБОТКИ 7](#_Toc197998583)

[3.1 Функциональное назначение 7](#_Toc197998584)

[3.2 Эксплуатационное назначение 7](#_Toc197998585)

[4. ТРЕБОВАНИЯ К ПРОГРАММЕ 8](#_Toc197998586)

[4.1 Требования к функциональным характеристикам 8](#_Toc197998587)

[4.1.1. Требования к составу выполняемых функций 8](#_Toc197998588)

[4.1.2. Требования к организации входных данных 8](#_Toc197998589)

[4.1.3. Требования к организации выходных данных 8](#_Toc197998590)

[4.1.4. Требования к временным характеристикам 8](#_Toc197998591)

[4.1.5. Требования к интерфейсу 9](#_Toc197998592)

[4.2 Требования к надежности 9](#_Toc197998593)

[4.3 Условия эксплуатации 9](#_Toc197998594)

[4.4 Требования к составу и параметрам технических средств 9](#_Toc197998595)

[4.5 Требования к информационной и программной совместимости 9](#_Toc197998596)

[4.6 Требования к маркировке и упаковке 9](#_Toc197998597)

[4.7 Требования к транспортировке и хранению 10](#_Toc197998598)

[5. ТРЕБОВАНИЯ К ПРОГРАММНОЙ ДОКУМЕНТАЦИИ 11](#_Toc197998599)

[6. ТЕХНИКО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ 12](#_Toc197998600)

[6.1 Ориентировочная экономическая эффективность 12](#_Toc197998601)

[6.2 Предполагаемая потребность 12](#_Toc197998602)

[6.3 Экономические преимущества разработки по сравнению с отечественными и зарубежными аналогами 12](#_Toc197998603)

[7. СТАДИИ И ЭТАПЫ РАЗРАБОТКИ 18](#_Toc197998604)

[ПОРЯДОК КОНТРОЛЯ И ПРИЕМКИ 19](#_Toc197998605)

[СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ 20](#_Toc197998606)

[Лист регистрации изменений 21](#_Toc197998607)

# АННОТАЦИЯ

Техническое задание – это основной документ, оговаривающий набор требований и порядок создания программного продукта, в соответствии с которым производится разработка программы, ее тестирование и приемка.

Настоящее Техническое задание на разработку «Приложение для разделения аудио на основе нейронных сетей» содержит следующие разделы: «Введение», «Основания для разработки», «Назначение разработки», «Требования к программе», «Требования к программным документам»,

«Технико-экономические показатели», «Стадии и этапы разработки», «Порядок контроля и приемки» и приложения.

В разделе «Введение» указано наименование, краткая характеристика и области назначения программы.

В разделе «Основания для разработки» указан документ, на основании которого ведется разработка, наименование темы разработки и ее условное обозначение.

В разделе «Назначение разработки» указано функциональное и эксплуатационное назначение программного продукта.

Раздел «Требования к программе» содержит основные требования к функциональным характеристикам, к надежности, к условиям эксплуатации, к составу технических и программных средств, к информационной и программной совместимости, к маркировке и упаковке, к транспортировке и хранению, а также специальные требования.

Раздел «Требования к программным документам» содержит предварительный состав программной документации и специальные требования к ней.

Раздел «Технико-экономические показатели» содержит ориентировочную экономическую эффективность, предполагаемую годовую потребность, экономические преимущества разработки программы.

Раздел «Стадии и этапы разработки» содержит стадии разработки, этапы и содержание работ.

В разделе «Порядок контроля и приемки» указаны общие требования к приемке работы.

Настоящий документ разработан в соответствии с требованиями:

1. ГОСТ 19.101-77 Виды программ и программных документов [1];
2. ГОСТ 19.102-77 Стадии разработки [2];
3. ГОСТ 19.103-77 Обозначения программ и программных документов [3];
4. ГОСТ 19.105-78 Общие требования к программным документам [5];
5. ГОСТ 19.106-78 Требования к программным документам, выполненным печатным способом [6];
6. ГОСТ 19.201-78 Техническое задание. Требования к содержанию и оформлению [7].

# ГЛОССАРИЙ

|  |  |
| --- | --- |
| **Термин** | **Определение** |
| Разделение аудио | Разделение аудиофайла с записанной музыкой на несколько дорожек с инструментами. |
| drag-and-drop | Это интерфейсная техника, позволяющая пользователю перемещать элементы на экране с помощью мыши или сенсорного ввода. Пользователь "захватывает" элемент, перетаскивает его в нужное место и "отпускает". В данном приложении эта техника должна поддерживаться для звуковых файлов. |
| solo/mute | Функции, используемые в аудиообработке и звуковом монтаже для изоляции или отключения отдельных дорожек. Solo - функция, которая позволяет прослушивать только выбранные дорожки, временно отключая все невыбранные. Mute - функция, которая полностью отключает звук на выбранной дорожке или источнике, позволяя сосредоточиться на других элементах микса без удаления дорожки. |
| waveform | Графическое представление звукового сигнала, отображающее изменение амплитуды звука во времени. |

# ВВЕДЕНИЕ

## Наименование программы

Наименование программы: «Приложение для разделения аудио на основе нейронных сетей».

Наименование программы на английским языке: «Application for audio separation based on neural networks».

Краткое наименование программы: «AudSep».

## Краткая характеристика области примения программы:

«AudSep» - это десктопное приложение, которое позволяет разделять аудиозаписи на отдельные инструментальные и вокальные дорожки с помощью технологии нейронных сетей. Данное приложение будет показывать высокое качество разделения благодаря использованию нейросетевых технологий.

# ОСНОВАНИЯ ДЛЯ РАЗРАБОТКИ

## Документы, на основании которых ведётся разработка

Разработка ведется на основании учебного плана подготовки бакалавров по направлению 09.03.04 “Программная инженерия” Факультета компьютерных наук Национального исследовательского университета “Высшая школа экономики” и утвержденной академическим руководителем программы темы курсового проекта.

## Наименование темы разработки

Наименование темы разработки: «Приложение для разделения аудио на основе нейронных сетей».

Наименование темы разработки на английским языке: «Application for audio separation based on neural networks».

Условное обозначение темы разработки: «AudSep».

# НАЗНАЧЕНИЕ РАЗРАБОТКИ

## Функциональное назначение

«AudSep» представляет собой программное обеспечение, предназначенное для автоматического разделения аудиозаписей на отдельные компоненты с использованием технологий машинного обучения.

Программа должна обеспечивать загрузку аудиофайлов через простой интерфейс с поддержкой drag-and-drop для звуковых файлов. Поддерживаются основные аудиоформаты включая MP3, WAV.

Основной функционал программы - разделение аудио на вокал и инструменты с помощью нейронной сети. Пользователь может выбрать предобученную модель. Во время обработки отображается прогресс-бар с возможностью отмены процесса.

Перед сохранением результатов пользователь может прослушать разделенные дорожки через встроенный плеер. Доступны базовые функции воспроизведения, регулировка громкости и выбор участка для прослушивания.

При экспорте пользователь выбирает формат вывода, качество и папку для сохранения. Программа автоматически именует файлы и может обрабатывать несколько треков в очереди. Все пользовательские настройки сохраняются между сессиями.

## Эксплуатационное назначение

Программный продукт рассчитан на пользователей различного уровня подготовки в области обработки аудио: от начинающих энтузиастов до профессиональных звукорежиссеров. Интуитивно понятный интерфейс не требует специальных навыков для базового использования программы.

Результаты работы программы могут быть использованы для дальнейшей обработки в других аудиоредакторах или как конечный продукт для создания караоке-версий, ремиксов и изоляции отдельных инструментов.

# ТРЕБОВАНИЯ К ПРОГРАММЕ

## Требования к функциональным характеристикам

### Требования к составу выполняемых функций

* Загрузка аудиофайлов форматов: MP3, WAV, FLAC.
* Разделение аудио на отдельные инструменты/вокал с помощью  нейросетевых моделей (HTDemucs, MelBand Roformer, BS Roformer).
* Предварительное прослушивание разделённых дорожек через встроенный аудиоплеер.
* Сохранение результатов разделения в файл на компьютере.
* Возможность выбрать нужные пользователю дорожки перед   
  сохранением.
* Отображение прогресса обработки.
* Регулировка громкости, solo/mute дорожек в плеере.
* Возможность настройки громкости отдельных дорожек перед экспортом в файл.
* Визуализация формы волны (waveform) для каждой дорожки.
* Возможность отмены текущей операции.

### Требования к организации входных данных

* Поддерживаемые форматы: MP3, WAV, FLAC.
* Максимальный размер файла: 100 МБ.
* Минимальное качество входного файла: 128 кбит/c (битрейт).
* Частота сэмплирования не менее, чем 44100 Гц.
* Поддержка моно и стерео файлов.

### Требования к организации выходных данных

* Форматы сохранения: WAV, MP3, FLAC.
* Возможность выбора инструментов для сохранения.
* Возможность регулировки громкости инструментов.

### Требования к временным характеристикам

* Время загрузки приложения: не более 3 секунд.
* Время обработки: не более, чем пятикратная длительность входного трека.
* Время отклика интерфейса: не более 100 мс.

### Требования к интерфейсу

* Возможность выбора места загрузки файла.
* Поддержка drag-and-drop для звуковых файлов.
* Визуализация прогресса обработки.

## Требования к надежности

* Должна выполняться проверка целостности входных файлов.
* При любых действиях пользователя программа не должна завершаться аварийно.
* Обработка ошибок с понятными сообщениями.
* Возможность отмены текущей операции.

## Условия эксплуатации

* Работа в операционных системах macOS 11+ .
* Наличие графического интерфейса.
* Доступ к файловой системе.

## Требования к составу и параметрам технических средств

Для надежной и бесперебойной работы программы требуется:

* ПК.
* Операционная система MacOS версии не ниже 11
* 8 гб оперативной памяти.

## Требования к информационной и программной совместимости

* Совместимость с macOS 11+.
* Программа должна быть написана на Python с использованием PyTorch.
* Программа должна использовать Python 3.8 или выше.

## Требования к маркировке и упаковке

Специальных требований к маркировке и упаковке не предъявляется. Программа должна распространяться в виде исполняемого файла или установочного файла.

## Требования к транспортировке и хранению

Специальных требований к транспортированию и хранению программного продукта не предъявляется: программа может передаваться и храниться на внешнем носителе (USB, CD, SSD, HDD или др.), в облачном хранилище в сети Интернет или с помощью любого другого способа хранения и передачи информации.

# ТРЕБОВАНИЯ К ПРОГРАММНОЙ ДОКУМЕНТАЦИИ

Предварительный состав программной документации:

1. Приложение для разделения аудио на основе нейронных сетей. Техническое задание (ГОСТ 19.201-78)
2. Приложение для разделения аудио на основе нейронных сетей. Программа и методика испытаний (ГОСТ 19.301-78)
3. Приложение для разделения аудио на основе нейронных сетей. Пояснительная записка (ГОСТ 19.404-79)
4. Приложение для разделения аудио на основе нейронных сетей. Руководство оператора (ГОСТ 19.505-79)
5. Приложение для разделения аудио на основе нейронных сетей. Текст программы (ГОСТ 19.401-78)

# ТЕХНИКО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ

## Ориентировочная экономическая эффективность

В рамках проекта расчёт экономической эффективности программного продукта не производился.

## Предполагаемая потребность

«AudSep» может использоваться музыкантами, звукорежиссерами или продюсерами для извлечения отдельных инструментальных дорожек из миксов, чтобы ремиксировать или анализировать композиции. Для улучшения качества звука, удаления нежелательных шумов или изменения баланса между различными элементами аудиозаписи.

## Экономические преимущества разработки по сравнению с отечественными и зарубежными аналогами

**Таблица 1.**

|  |  |  |
| --- | --- | --- |
| **Название** | **Ссылка** | **Описание** |
| AudSep | https://github.com/prikokes/AudSep | Десктоп приложение для разделения аудио на музыкальные инструменты. Поддерживается WAV, MP3 форматы файлов. |
| MoisesAI | https://moises.ai/ru/ | MoisesAI - это онлайн-платформа, использующая искусственный интеллект для разделения музыкальных треков на отдельные инструменты и вокал. Сервис позволяет извлекать из готовых композиций вокал, бас, ударные, гитару и другие инструменты, а также предлагает дополнительные функции, такие как определение аккордов и изменение темпа или тональности.  Платформа доступна через веб-интерфейс и мобильные приложения, поддерживает форматы MP3 и WAV. Сервис работает по модели подписки: есть бесплатная версия с ограничениями и платные тарифы с расширенными возможностями.  MoisesAI популярен среди музыкантов для разбора партий, создания минусовок и обучения, а также среди продюсеров для работы с ремиксами. |
| SplitterAI | https://vocalremover.org/ru/splitter-ai | SplitterAI - это бесплатный онлайн-сервис для разделения аудиотреков на составляющие компоненты с использованием искусственного интеллекта. В отличие от многих конкурентов, сервис не требует регистрации и предлагает базовый функционал без платной подписки.  Сервис позволяет разделить музыкальную композицию на четыре основные дорожки: вокал, ударные, бас и остальные инструменты. Работает с файлами формата MP3, WAV и других популярных аудиоформатов. Максимальный размер загружаемого файла ограничен 50 МБ.  Основные особенности SplitterAI - это простой минималистичный интерфейс, быстрая обработка файлов и отсутствие необходимости установки дополнительного программного обеспечения. Все операции выполняются через веб-браузер. После обработки пользователь может скачать как отдельные дорожки, так и их комбинации. |
| LALAL.AI | https://www.lalal.ai/ | LALAL.AI - это онлайн-платформа для разделения аудио, использующая собственную технологию Phoenix на основе нейронных сетей. Главное отличие сервиса - высокое качество разделения и возможность работы не только с аудио, но и с видеофайлами.  Сервис предлагает несколько режимов разделения: стандартный Phoenix и улучшенный Phoenix Stems Pro. Пользователи могут извлекать вокал, инструменты, ударные, бас и другие компоненты из аудиотреков. При работе с видео LALAL.AI сохраняет исходное качество видеоряда, заменяя только аудиодорожку.  Поддерживаются все популярные форматы: MP3, WAV, FLAC, MP4, AVI, MOV и другие. Размер загружаемых файлов может достигать 2 ГБ (зависит от тарифа). Обработка происходит на серверах компании с сохранением конфиденциальности - файлы удаляются после завершения работы.  Сервис работает по системе покупки минут обработки, а не по подписке. Новым пользователям доступны бесплатные пробные минуты для тестирования. Интерфейс простой и интуитивно понятный, доступен на разных языках, включая русский. |
| RipX DAW | https://hitnmix.com/ripx-daw/ | RipX DAW - это профессиональное десктопное приложение от компании Audionamix для разделения и редактирования аудио. Это полноценная рабочая станция (DAW), специализирующаяся на работе с отдельными компонентами аудиотреков.  Программа использует продвинутые алгоритмы искусственного интеллекта для разделения музыки на составляющие: вокал, бас, ударные, гитару и другие инструменты. Особенность RipX - возможность глубокого редактирования каждой выделенной дорожки, включая точную настройку частот, изменение тональности и темпа.  Ключевые возможности включают редактирование отдельных нот, изменение мелодических линий, удаление или добавление инструментов, коррекцию вокала и создание ремиксов. Программа также предлагает инструменты для спектрального редактирования и очистки аудио от шумов.  RipX DAW доступен в двух версиях: стандартной и профессиональной. Профессиональная версия включает дополнительные инструменты для редактирования и более точные алгоритмы разделения. Программа требует мощный компьютер для обработки, но работает автономно, без необходимости подключения к интернету. |

# СТАДИИ И ЭТАПЫ РАЗРАБОТКИ

|  |  |  |  |
| --- | --- | --- | --- |
| Стадия  разработки | Этап работ | Содержание работ | Срок  выполнения |
| 1. Техническое задание | Обоснование необходимости разработки программы | * Постановка задачи * Сбор исходных материалов | 28 октября  2024 – 10 января 2025 |
| Научно-  исследовательские работы | * Определние структуры входны и выходных данных * Предварительный выбор методов решения задач. |
| Разработка и утверждение технического задания | * Определение стадий, этапов и сроков разработки программы и документации на нее * Выбор языков программирования. * Согласование и утверждение технического задания. |
| 2. Рабочий проект | Разработка программы | * Рарзработка и отладка программы * Разработка программной документации | 10 10 января  2025 – 1 марта 2025 |
| Испытание программы | * Разработка, согласование и утверждение порядка и методики испытаний. * Корректировка программы и программной документации по результатам испытаний. |
| 3. Внедрение | Разработка программной документации | * Разработка программных документов в соответствии с требованиями ГОСТ 19 ЕСПД * Выверка и корректировка неточностей, допущенных при разработке программной документации. | 1 марта 2025 – 30 марта 2025 |
| 454. Защита проекта комиссии | | * Формирование архива, содержащего разработанную программу и документацию к ней, и его загрузка в систему SmartLMS НИУ ВШЭ. * - Подготовка презентации и защита проекта комиссии. | Май 2025 |

Исполнитель: Иванов Георгий, студент БПИ-223

# ПОРЯДОК КОНТРОЛЯ И ПРИЕМКИ

Контроль и приемка разработки осуществляются в соответствии с документом «Программа и методика испытаний» (ГОСТ 19.301–79).

Защита выполненного проекта осуществляется комиссии, состоящей из преподавателей департамента программной инженерии, в утверждённые приказом декана ФКН сроки.

# СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

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 19.201-78 Техническое задание. Требования к содержанию и оформлению. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
8. ГОСТ 19.603-78 Общие правила внесения изменений. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
9. ГОСТ 19.604-78 Правила внесения изменений в программные документы, выполненные печатным способом. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
10. ГОСТ 15150-69 Машины, приборы и другие технические изделия. Исполнения для различных климатических районов. Категории, условия эксплуатации, хранения и транспортирования в части воздействия климатических факторов внешней среды. – М.: Изд-во стандартов, 1997.
11. ГОСТ 19.301-79 Программа и методика испытаний. Требования к содержанию и оформлению. //Единая система программной документации. – М.: ИПК Издательство стандартов, 2001.
12. Rouard S., Massa F., Défossez A. Hybrid transformers for music source separation //ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – IEEE, 2023. – С. 1-5.
13. Lu W. T. et al. Music source separation with band-split rope transformer //ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – IEEE, 2024. – С. 481-485.
14. Wang J. C., Lu W. T., Won M. Mel-Band RoFormer for Music Source Separation //arXiv preprint arXiv:2310.01809. – 2023.
15. ZFTurbo. Music-Source-Separation-Training [Электронный ресурс]. – Режим доступа: https://github.com/ZFTurbo/Music-Source-Separation-Training (дата обращения: 01.04.2025).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Лист регистрации изменений | | | | | | | | | |
| Номера листов (страниц) | | | | | Всего листов (страниц в докум.) | № документа | Входящий  № сопроводит ельного докум. и дата | Подп. | Дата |
| Изм. | Измененных | Замененных | Новых | Аннулированн ых |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |