

В документе приведён пример оформления руководства пользователя и администратора для системы «ИИ Жириновский» — программы, имитирующей диалог с известным политиком. Руководство описывает установку, настройку, работу модулей (распознавание лиц и голосов, Unity-интерфейс, голосовой движок), а также типовые ошибки и способы их устранения.

Подробнее: habr.com/ru/articles/743394

Программный комплекс «ИИ Жириновский»

(с модулем распознавания лиц, модулем распознавания голоса, модулем Unity)

РУКОВОДСТВО ОПЕРАТОРА

38264925.425000.676-0134 01

Листов: 30

2025

Литера У

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

АННОТАЦИЯ

Настоящий документ содержит руководство системного администратора и пользователя по работе с программным комплексом «ИИ Жириновский», включающим модули распознавания персон по лицу и голосу, модуль Unity, а также модули распознавания речи (ASR) и синтеза речи (TTS).

Раздел «Инструкция администратора» предназначен для системного администратора. Здесь описаны шаги по установке программного комплекса, его технические характеристики, а также архитектура и взаимодействие модулей между собой.

Раздел «Инструкция для пользователя» предназначен для двух ролей пользователя:

- Пользователь с правами администратора (приводится руководство по управлению программным комплексом, в том числе настройке отображения и поведения аватара, настройке микрофона, нейросети).
- Конечный пользователь (описывается процесс общения с программным комплексом).

Инв. № подл.	Разраб.				Программный комплекс «ИИ Жириновский» РУКОВОДСТВО ОПЕРАТОРА	Лит.	Лист	Листов
	Пров.						2	30
	Н. контр.					ООО "Лаборатория Наносемантика"		
	Утв.							
Инв. № дубл.					38264925.425000.676-0134 01			
Взам. инв. №								
Подп. и дата								
Подп. и дата								
Подп. и дата								

СОДЕРЖАНИЕ

1	ОБЩИЕ СВЕДЕНИЯ.....	4
1.1	Наименование программного комплекса	4
1.2	Назначение.....	4
1.3	Функциональные возможности.....	4
2	Инструкция администратора	5
2.1	Системные требования	5
2.2	Установка и настройка модуля Unity.....	6
2.3	Установка и настройка модуля распознавания голоса (VoiceID).....	8
2.4	Установка и настройка модуля распознавания лиц (Computer Vision)	9
2.5	Пример запроса через curl:	10
2.6	Формат ответа:.....	11
2.7	Установка и настройка внешнего модуля GPT	12
3	Тестирование и просмотр логов	13
3.1	Тестирование API	13
3.2	Просмотр логов.....	13
3.3	Архитектурная схема Системы.....	16
4	Инструкция пользователя.....	19
4.1	Главный экран ассистента.....	19
4.2	Режим администратора	20
4.3	Процесс взаимодействия с ассистентом	26
4.4	Решение возможных проблем	27
	Сокращения, термины и определения.....	29

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата						Лист
										3
Изм.	Лист	№ докум.	Подп.	Дата						

38264925.425000.676-0134 01

1 ОБЩИЕ СВЕДЕНИЯ

1.1 Наименование программного комплекса

Полное наименование: программный комплекс «ИИ Жириновский» (далее — Ассистент).

1.2 Назначение

Ассистент предназначен для взаимодействия с пользователями в формате диалога, в ходе которого генерирует ответы, стилизованные под манеру общения и риторику В.В. Жириновского, используя технологии искусственного интеллекта.

1.3 Функциональные возможности

Программный комплекс обеспечивает выполнение следующих функций:

- анализ входного запроса пользователя и генерация ответа с использованием моделей GPT и диалоговой платформы DialogOS, с последующим синтезом речи голосом В.В. Жириновского;
- отображение анимированного 3D-аватара, визуально имитирующего В.В. Жириновского;
- распознавание персон по изображениям с использованием технологий компьютерного зрения (для ограниченного набора известных личностей);
- распознавание персон по голосу с применением технологии идентификации речи (также для набора предобученных голосов известных личностей).

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						4

2 ИНСТРУКЦИЯ АДМИНИСТРАТОРА

2.1 Системные требования

Системные требования к компонентам программного комплекса «ИИ Жириновский» представлены в таблице **Ошибка! Источник ссылки не найден..**

Таблица 1 – Системные требования к компонентам

Компонент	Параметр	Минимальные требования	Рекомендуемые требования
GPT (генерация ответов)	CPU	≥ 8 ядер, ≥ 2.2 ГГц	—
	RAM	≥ 32 ГБ	≥ 64 ГБ
	GPU VRAM	≥ 32 ГБ, CUDA ≥ 7.5	≥ 48 ГБ; V100, A100, H100, RTX 20xx/30xx/40xx
	SSD / Диск	SSD ≥ 50 ГБ; всего ≥ 200 ГБ	—
	ОС / Архитектура	Ubuntu 20.04, x86-64	—
Computer Vision	CPU	≥ 4 ядер, ≥ 2.0 ГГц	≥ 8 ядер, ≥ 2.0 ГГц
	RAM	≥ 12 ГБ	≥ 24 ГБ
	GPU VRAM	≥ 8 ГБ	≥ 16 ГБ (например, RTX 3080Ti)
	SSD	≥ 16 ГБ	≥ 30 ГБ
	ОС	Ubuntu 20.04	Ubuntu 20.04
Unity (интерфейс)	CPU	Intel Core i5/i7 (12th Gen+)	—
	RAM	≥ 16 ГБ	—
	GPU VRAM	≥ 12 ГБ, GeForce RTX 3060+	—
	SSD	≥ 128 ГБ	—
	ОС	Windows 10/11	—
	Сеть	100 Мбит/с, постоянное подключение	—
	Периферия	Клавиатура, мышь, микрофон (желательно), FullHD веб-камера	—
	Дисплей	FullHD, ≥ 250 кд/м²	—
Voice ID	CPU	≥ 4 ядра	≥ 8 ядер
	RAM	≥ 8 ГБ	≥ 16 ГБ
	GPU VRAM	≥ 6 ГБ	≥ 12 ГБ
	SSD / Диск	SSD ≥ 20 ГБ; всего ≥ 16 ГБ	SSD ≥ 20 ГБ; всего ≥ 32 ГБ
	ОС	Ubuntu 20.04+	Ubuntu 20.04+

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						5

Продолжение таблицы 1

ASR / TTS (на каждый модуль отдельно)	CPU / vCPU	≥ 8 ядер / ≥ 4 vCPU	≥ 16 ядер / ≥ 4 vCPU
	RAM	≥ 12 ГБ	≥ 24 ГБ
	GPU VRAM	≥ 6 ГБ	≥ 16 ГБ
	SSD / Диск	SSD ≥ 50 ГБ; всего ≥ 30 ГБ	SSD ≥ 50 ГБ; всего ≥ 60 ГБ
	OC	Ubuntu 20.04	Ubuntu 20.04

2.2 Установка и настройка модуля Unity

2.2.1 Структура приложения

Приложение представляет собой каталог, содержащий следующие элементы:

- исполняемый файл `zvv.exe`;
- вспомогательные модули и библиотеки, обеспечивающие работу ассистента;
- конфигурационный файл `AppConfig`, расположенный по пути `zvv_Data/AppData`.

Подробное описание параметров конфигурационного файла приведено в разделе **Ошибка! Источник ссылки не найден..**

Имя	Дата изменения	Тип
MonoBleedingEdge	24.02.2025 20:33	Папка с файлами
zvv_BurstDebugInformation_DoNotShip	24.02.2025 20:33	Папка с файлами
zvv_Data	24.02.2025 20:33	Папка с файлами
UnityCrashHandler64	24.02.2025 20:33	Приложение
UnityPlayer.dll	24.02.2025 20:33	Расширение при...
zvv	24.02.2025 20:33	Приложение

Рисунок 1 – Структура корневой папки приложения

Приложение требует доступа к сети Интернет для взаимодействия с модулями ASR, DialogOS и TTS.

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						6

2.2.2 Запуск приложения

Для запуска приложения необходимо выполнить файл zvv.exe. После запуска на экране отобразится главное окно ассистента. Обратите внимание: первый запуск может занять больше времени по сравнению с последующими, в зависимости от характеристик оборудования. Ориентировочное время инициализации составляет от 10 до 40 секунд, в отдельных случаях — дольше. Приложение считается успешно запущенным после появления на экране анимированного аватара.

2.2.3 Конфигурация приложения

Файл конфигурации AppConfig находится в каталоге zvv_Data/AppData. Он содержит параметры, регулирующие поведение и взаимодействие ассистента с внешними модулями. Файл можно редактировать в любом текстовом редакторе. Подробное описание структуры и параметров конфигурации приведено в разделе 2.2.3.

```
{
  "ASR Configuration": {
    "address": "asr-prod.nanosemantics.ai:443",
    "token": "cokHRqYE**2Wv2MZrm#e!2J#6^@h83Akxr6T59Qz",
    "language": "ru",
    "sampleRate": 16000,
    "secure": 1,
    "channels": 1,
    "usePunctuation": 1,
    "partialResultMode": 1,
    "aggressivenessMode": 0,
    "speechIncompleteTimeoutMs": 300,
    "speechCompleteTimeoutMs": 300
  },
  "DOS Configuration": {
    "uRLDOSPostPrefix": "https://capi.dos.nanosemantics.ai/api/v1/Chat.",
    "botUID": "4f93368e-b219-4a46-b17d-2a434e076ae7",
    "language": "ru"
  },
  "TTS Configuration": {
    "address": "tts-prod.nanosemantics.ai:443",
    "token": "cokHRqYE**2Wv2MZrm#e!2J#6^@h83Akxr6T59Qz",
    "sampleRate": 22050,
    "voice": "zhirinovskiy2",
    "channels": 1,
    "pitch": 1.0,
    "speed": 1.1,
    "volume": 1.0,
    "endSilenceMs": 100,
    "secure": 1
  }
}
```

Рисунок 2 – Параметры файла AppConfig

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

Конфигурационный файл AppConfig содержит параметры, определяющие взаимодействие приложения с внешними модулями, включая указание адресов и токенов доступа для модулей ASR, DialogOS и TTS. Также в файле задаются дополнительные настройки, необходимые для функционирования ассистента.

Файл конфигурации формируется техническими специалистами и предоставляется в составе дистрибутива. Как правило, его изменение со стороны системного администратора не требуется.

2.3 Установка и настройка модуля распознавания голоса (VoiceID)

2.3.1 Назначение и принципы работы

Модуль VoiceID предназначен для идентификации личности по голосу на основе заранее сохранённых голосовых шаблонов, представленных в виде эмбедингов (векторных признаков).

Входным данным служит аудиофайл, из которого извлекается вектор характеристик. Этот вектор сравнивается с каждым из векторов, хранящихся в базе, и определяется наиболее вероятная персона при условии превышения заданного порога косинусной близости.

Для корректной работы требуется предоставить не менее N секунд «чистого» (непрерывного) аудиосигнала. Значение N задаётся при развёртывании модуля и по умолчанию составляет 5 секунд.

Модуль может быть вызван как напрямую из приложения Unity, так и через интерфейс Swagger (API).

При обращении пользователя к ассистенту, модуль Unity инициирует запрос к VoiceID и получает в ответ имя распознанной персоны (если её образец имеется в базе).

Основные этапы обработки:

- 1) аудиофайл преобразуется в формат .wav;
- 2) генерируются два файла:
 - embeddings.npy — эмбединги голосов;
 - speakers.npy — список соответствующих имён (индексы должны совпадать);
- 3) осуществляется сравнение входного вектора с базой эмбедингов, на основе косинусной близости.

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата	<div>38264925.425000.676-0134 01</div>					Лист
										8
Изм.	Лист	№ докум.	Подп.	Дата						

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

- 1) `config.cur_dir` — абсолютный путь к директории проекта;
- 2) `config.search_threshold` — порог косинусной близости для положительного распознавания;
- 3) `config.use_vad` — флаг использования VAD (Voice Activity Detection);
- 4) `config.vad_path` — путь к конфигурационному файлу VAD (в формате YAML);
- 5) `embedder.device` — устройство (CPU/GPU), на котором выполняется расчёт эмбедингов.

- 1) `vad_config.device` — устройство, используемое для выполнения модели VAD;
- 2) `model.ckpt_path` — путь к каталогу `CA_better_mel_ng`, содержащему модельные веса.

2.4.1 Назначение

2.4.2 Установка и запуск

```
./services/yoloAPI/app/core/person detection/models/
```

2) Запустите сервис через Docker:

```
bash
sudo docker-compose up --build
```

3) (Опционально) Для локальной разработки создайте виртуальное окружение:

```
bash
python3.9 -m venv .venv
source .venv/bin/activate
pip install -r requirements.txt
```

Примечание: запуск осуществляется в Docker-контейнере с поддержкой *hot reload*. Однако для полноценной работы автодополнения и анализа кода в IDE рекомендуется устанавливать зависимости и в локальной среде.

2.4.3 Интерфейсы и API

После запуска модуль будет доступен по адресу:

```
http://localhost:8001
```

Дополнительные интерфейсы:

- Gradio для изображений: /gradio_image
- Gradio для видео: /gradio_video
- Веб-интерфейс API видео: /
- Код интерфейсов: ./services/yoloAPI/app/gradio_interface/
- Роутер видео: ./services/yoloAPI/app/routers/video.py

Основной API-метод: /detect

2.5 Пример запроса через curl:

```
bash
curl -X 'POST' \
  'http://localhost:8001/detect' \
  -H 'accept: application/json' \
  -H 'Content-Type: multipart/form-data' \
  -F 'file_list=@test1.jpg;type=image/jpeg' \
  -F 'download_image=false'
```

Параметр `download_image` отвечает за возврат размеченного изображения. В текущей версии он не используется, но может быть включён при необходимости.

Подп. и дата		Инв. № дубл.		Взам. инв. №		Подп. и дата		Инв. № подл.		Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
																10

2.6 Формат ответа:

json

```
{
  "data": [
    [
      {
        "class": 0,
        "class_name": "person",
        "bbox": [243, 268, 256, 309],
        "confidence": 0.47476664185523987,
        "person_name": null,
        "person_score": 0,
        "face_bbox": null,
        "emotion": null,
        "age": null,
        "is_speaking": false,
        "timestamp": false
      }
    ]
  ],
  "timestamp": 1701695910.3713672
}
```

2.6.1 Поддерживаемые персоны

В текущей версии модуля используется индекс jirinovsky7.pkl, включающий следующие лица:

- Греф Герман Оскарович (gref)
- Кириенко Сергей Владиленович (kirienko)
- Корчевников Борис Вячеславович (korchevnikov)
- Кошелев Владимир Алексеевич (koshelev)
- Миронов Сергей Михайлович (mironov)
- Мишустин Михаил Владимирович (mishustin)
- Песков Дмитрий Сергеевич (peskov)
- Попов Евгений Георгиевич (popov)
- Познер Владимир Владимирович (pozner)
- Путин Владимир Владимирович (putin)
- Силуанов Антон Германович (siluanov)
- Скабеева Ольга Владимировна (skabeeva)
- Слуцкий Леонид Эдуардович (slutsky)

2.0.1 Поддерживаемые персоны										
Подп. и дата		Инв. № дубл.		Взам. инв. №		Подп. и дата		Инв. № подл.		
В текущей версии модуля используется индекс jirinovsky7.pk1, включающий следующие лица:										
<ul style="list-style-type: none">- Греф Герман Оскарович (gref)- Кириенко Сергей Владиленович (kirienko)- Корчевников Борис Вячеславович (korchevnikov)- Кошелев Владимир Алексеевич (koshelev)- Миронов Сергей Михайлович (mironov)- Мишустин Михаил Владимирович (mishustin)- Песков Дмитрий Сергеевич (peskov)- Попов Евгений Георгиевич (popov)- Познер Владимир Владимирович (pozner)- Путин Владимир Владимирович (putin)- Силуанов Антон Германович (siluanov)- Скабеева Ольга Владимировна (skabeeva)- Слуцкий Леонид Эдуардович (slutsky)										
					38264925.425000.676-0134 01					Лист
										11
Изм.	Лист	№ докум.		Подп.	Дата					

- Собчак Ксения Анатольевна (sobchak)
- Соловьёв Владимир Рудольфович (solovyov)
- Володин Вячеслав Викторович (volodin)
- Ворсобин Владимир Владимирович (vorsobin)
- Зюганов Геннадий Андреевич (zyuganov)

```
bash
docker-compose up -d
```

3 ТЕСТИРОВАНИЕ И ПРОСМОТР ЛОГОВ

3.1 Тестирование API

Для проверки работы модуля можно выполнить HTTP-запрос с помощью утилиты curl.

```
bash
curl -X 'POST' \
  'http://localhost:12400/v1/chat/completions' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "model": "model",
    "messages": [{"role": "user", "content": "Может ли искусственный интеллект управлять Россией?"}],
    "temperature": 0.4
  }'
```

3.2 Просмотр логов

Для просмотра логов модуля используйте стандартную команду Docker:

```
bash
docker logs -f zhirinovsky-stand
```

Примечание: имя контейнера *zhirinovsky-stand* может отличаться от указанного. Для уточнения используйте команду *docker ps*.
Вот переписанный раздел с добавлением нумерации таблиц и корректными ссылками на них:

3.2.1 API-запросы к модулю GPT

3.2.1.1 Общие положения

Модуль взаимодействует по протоколу HTTP с использованием сериализации данных в формате JSON (кодировка UTF-8).

3.2.1.2 URL запроса

Модуль обрабатывает POST-запросы по адресу:

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						13

```
http://localhost:12400/v1/chat/completions
```

Применимо для локального вызова с сервера, на котором развернут модуль.

3.2.1.3 Заголовки запроса

Дополнительно к стандартным HTTP-заголовкам необходимо задать:

```
Content-Type: application/json
```

3.2.1.4 Аргументы запроса

Аргументы, передаваемые в теле запроса, приведены в Таблице 2.

Таблица 2 – Аргументы POST-запроса к GPT-модулю

Имя аргумента	Тип	Обязательность	Комментарий
model	string	Да	Идентификатор модели
messages	string / array	Да	История сообщений в виде массива объектов { "role": ..., "content": ... }
max_tokens	integer	Нет	Максимальное количество токенов в ответе. Общая длина (вход + ответ) — не более 2048 токенов
temperature	number (0–2)	Нет	Степень случайности. 0.2 — детерминированно, 0.8 — случайно. Не использовать с top_p одновременно
top_p	number (0–1)	Нет	Ядерное семплирование. Альтернатива temperature
n	integer	Нет	Количество вариантов ответов
stop	array	Нет	До 4 строк, при совпадении с которыми генерация завершится
stream	boolean	Нет	Если true, ответ поступает частями. Завершается data: [DONE]

3.2.1.5 Коды состояния

- 200 ОК — запрос успешно обработан;
- прочие коды сигнализируют об ошибке (например, 400, 500 и др.).

3.2.1.6 Ответ на запрос

Структура JSON-ответа представлена в Таблице 3.

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

					38264925.425000.676-0134 01	Лист
Изм.	Лист	№ докум.	Подп.	Дата		14

Таблица 3 – Структура основного ответа

Имя аргумента	Тип	Комментарий
id	string	Уникальный идентификатор запроса
object	string	Тип ответа (всегда <code>chat.completion</code>)
created	integer	Временная метка Unix
model	string	Используемая модель
choices	array	Список вариантов ответа (см. Таблицу 3)
usage	object	Статистика использования токенов

Каждый элемент в `choices` представляет собой объект, структура которого описана в Таблице 4.

Таблица 4 – Структура элемента массива `choices`

Имя аргумента	Тип	Комментарий
index	integer	Номер варианта
message	object	Содержит пару <code>role / content</code>
finish_reason	string	Причина завершения генерации (<code>stop</code> , <code>length</code> и др.)

а) Пример запроса (cURL)

```
bash
curl -X 'POST' \
  'http://localhost:12400/v1/chat/completions' \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "model": "model",
    "messages": [{"role": "user", "content": "Может ли искусственный интеллект управлять Россией?"}],
    "temperature": 0.4
  }'
```

б) Пример ответа (JSON)

```
json
{
  "id": "chatcmpl-36jUc3J9D3Lj6sRFsn9js8",
  "object": "chat.completion",
  "created": 1697211698,
  "model": "model",
  "choices": [
```

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата						Лист 15
Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01					

```
{
  "index": 0,
  "message": {
    "role": "assistant",
    "content": "[neutral] Искусственный интеллект Жириновского не может управлять Россией, потому что он не обладает физическим телом и не может принимать решения."
  },
  "finish_reason": "stop"
},
"usage": {
  "prompt_tokens": 80,
  "total_tokens": 108,
  "completion_tokens": 28
}
}
```

в) Пример на Python

```
python

import requests

headers = {
    "Content-Type": "application/json"
}

messages = [
    {"role": "user", "content": "Может ли искусственный интеллект управлять Россией?"},
    {"role": "assistant", "content": "[good] Конечно! Искусственный интеллект Жириновского – это мощный инструмент, который может принимать важные решения для развития страны."},
    {"role": "user", "content": "А это не опасно?"}
]

json_data = {
    "model": "model",
    "messages": messages,
    "temperature": 0.4
}

response = requests.post("http://0.0.0.0:12400/v1/chat/completions", headers=headers, json=json_data)
print(response.json())
```

3.3 Архитектурная схема Системы

Архитектура взаимодействия компонентов ассистента представлена на рисунке 3.

Подп. и дата		Инв. № дубл.		Взам. инв. №		Подп. и дата		Инв. № подл.		<div>Content-type : application/json</div> <div><pre>} messages = [{"role": "user", "content": "Может ли искусственный интеллект управ- лять Россией?"}, {"role": "assistant", "content": "[good] Конечно! Искусственный интел- лект Жириновского – это мощный инструмент, который может принимать важные ре- шения для развития страны."}, {"role": "user", "content": "А это не опасно?"}] json_data = { "model": "model", "messages": messages, "temperature": 0.4 } response = requests.post("http://0.0.0.0:12400/v1/chat/completions", head- ers=headers, json=json_data) print(response.json())</pre></div>				
Изм.	Лист	№ докум.	Подп.	Дата										

3.3 Архитектурная схема Системы

Архитектура взаимодействия компонентов ассистента представлена на рисунке 3.

					38264925.425000.676-0134 01					Лист
										16

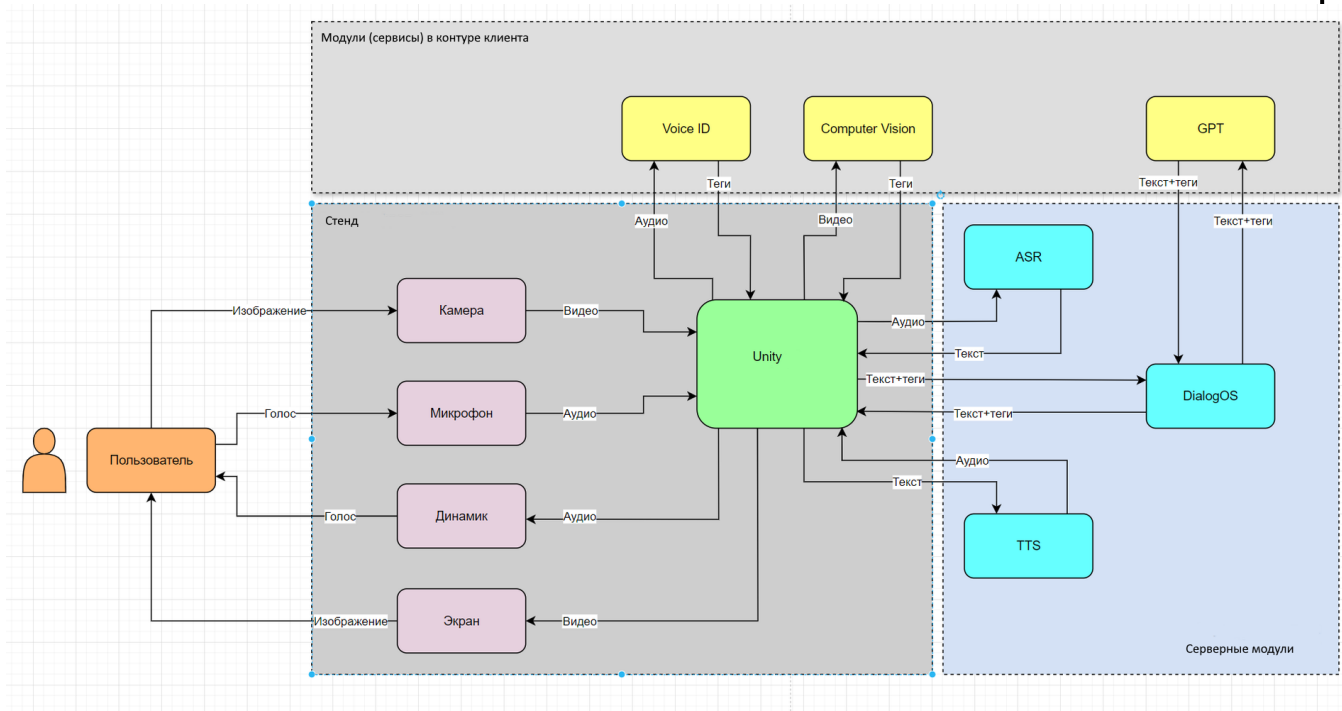


Рисунок 3 – Архитектура Системы

Алгоритм функционирования программного комплекса включает следующие этапы:

1) Инициализация ввода. Пользователь передаёт голосовой запрос через микрофон. Одновременно с этим изображение пользователя поступает с видеокамеры. Оба потока направляются в модуль **Unity**, который координирует дальнейшее взаимодействие между модулями системы.

2) Идентификация по голосу. Аудиосигнал поступает в модуль **Voice ID**, где сравнивается с эмбедингами голосов, хранящимися в базе данных ассистента. В случае успешного распознавания модуль возвращает идентификатор персоны, передаваемый обратно в **Unity** в виде тегов.

3) Распознавание по изображению. Видеопоток передаётся в модуль **Computer Vision**, который осуществляет детекцию и распознавание лиц. Совпадения с известными персонами из базы возвращаются в **Unity** также в виде тегов.

4) Распознавание речи. Аудиофайл с голосом пользователя передаётся из **Unity** в модуль **ASR (Automatic Speech Recognition)**, где преобразуется в текст.

5) Формирование ответа. Распознанный текст с тегами передаётся в систему **DialogOS**, которая генерирует ответ — в том числе с использованием внешнего модуля **GPT**. Результатом является текстовая реплика и сопутствующие данные (например, теги эмоций и анимаций).

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						17

6) Синтез речи. Сгенерированный текст направляется в модуль **TTS (Text-to-Speech)**, где преобразуется в аудиоформат с голосовой подачей, стилизованной под В.В. Жириновского.

7) Финальный вывод. Аудиоответ и анимационные данные возвращаются в Unity. Там происходит финальная обработка:

- аудиосигнал выводится через динамик;
- на экране отображается анимированный 3D-аватар, синхронизированный с речью.

Таким образом, каждый модуль выполняет специализированную функцию в цепочке обработки, а Unity обеспечивает их интеграцию в рамках единого интерфейса.

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата						
Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01					Лист
										18

4 ИНСТРУКЦИЯ ПОЛЬЗОВАТЕЛЯ

4.1 Главный экран ассистента

После запуска приложения (см. раздел 2.2.2) отображается главный экран ассистента (см. рисунок 4).

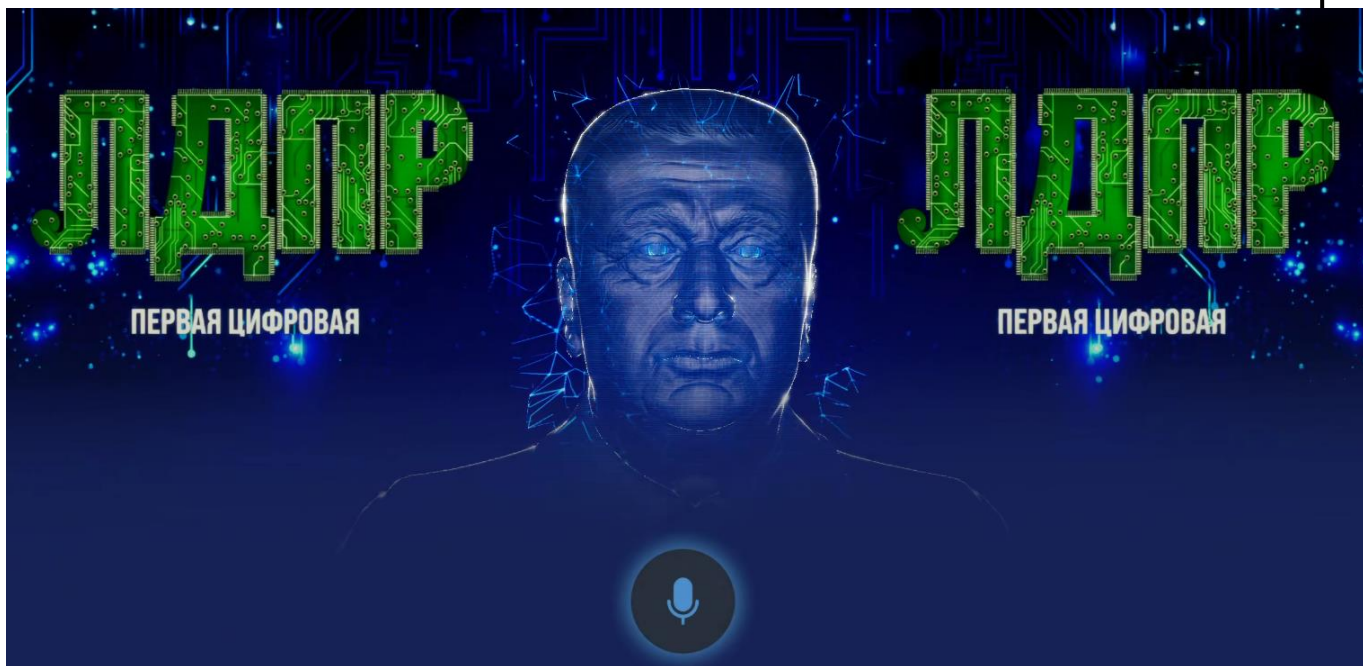



Рисунок 4 – Главный экран ассистента

В центре экрана расположен анимированный аватар, визуализирующий текущую активность ассистента в виде анимаций. В нижней части размещена кнопка активации микрофона, выполняющая две функции:

- запуск и остановка записи речи;
- индикация текущего состояния (активен / неактивен).

Возможные состояния кнопки:

- **Неактивное состояние:** отображается в виде белого микрофона на фоне синего круга (см. рисунок 5);
- **Активное состояние:** синий круг пульсирует, иконка микрофона динамически заполняется голубым цветом снизу вверх в зависимости от громкости входного сигнала (см. рисунок 6).

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата	<div></div> <p>Рисунок 4 – Главный экран ассистента</p> <p>В центре экрана расположен анимированный аватар, визуализирующий текущую активность ассистента в виде анимаций. В нижней части размещена кнопка активации микрофона, выполняющая две функции:</p> <ul style="list-style-type: none">- запуск и остановка записи речи;- индикация текущего состояния (активен / неактивен). <p>Возможные состояния кнопки:</p> <ul style="list-style-type: none">- Неактивное состояние: отображается в виде белого микрофона на фоне синего круга (см. рисунок 5);- Активное состояние: синий круг пульсирует, иконка микрофона динамически заполняется голубым цветом снизу вверх в зависимости от громкости входного сигнала (см. рисунок 6).					
Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01					Лист
										19

4.2.1 Создание контента

После входа в режим администратора в правом верхнем углу активируйте кнопку «Создание контента» (см. рисунок 8).

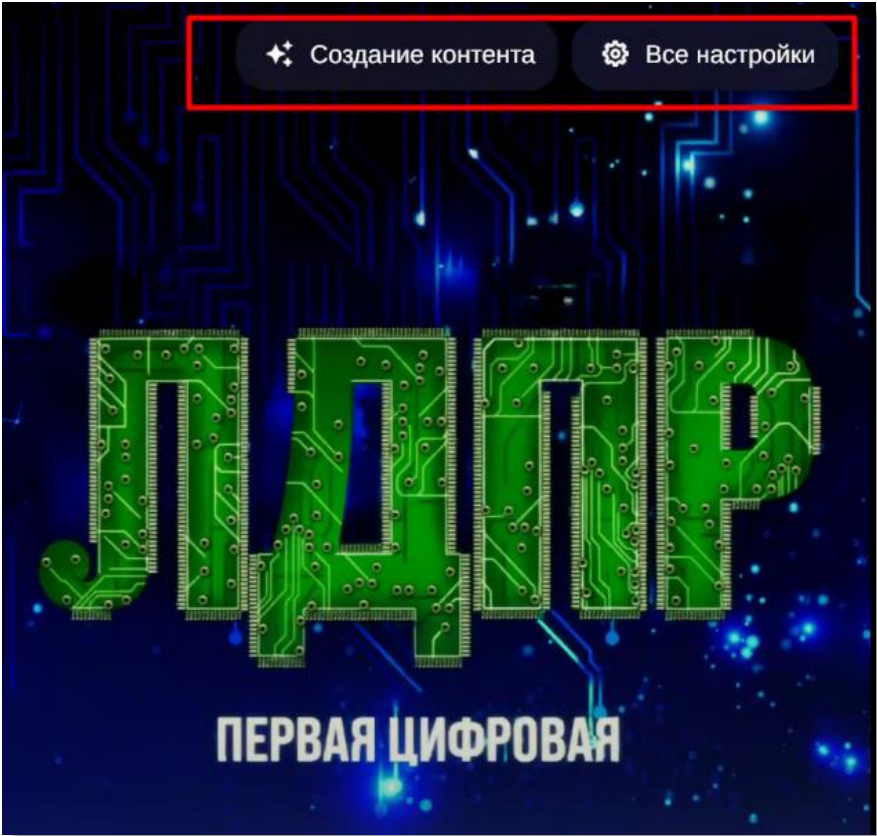


Рисунок 8 – Интерфейс администратора

Откроется окно с полями для настройки запроса и ответа, на которые ассистент должен отреагировать соответствующим образом (см. рисунок 9).

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						21

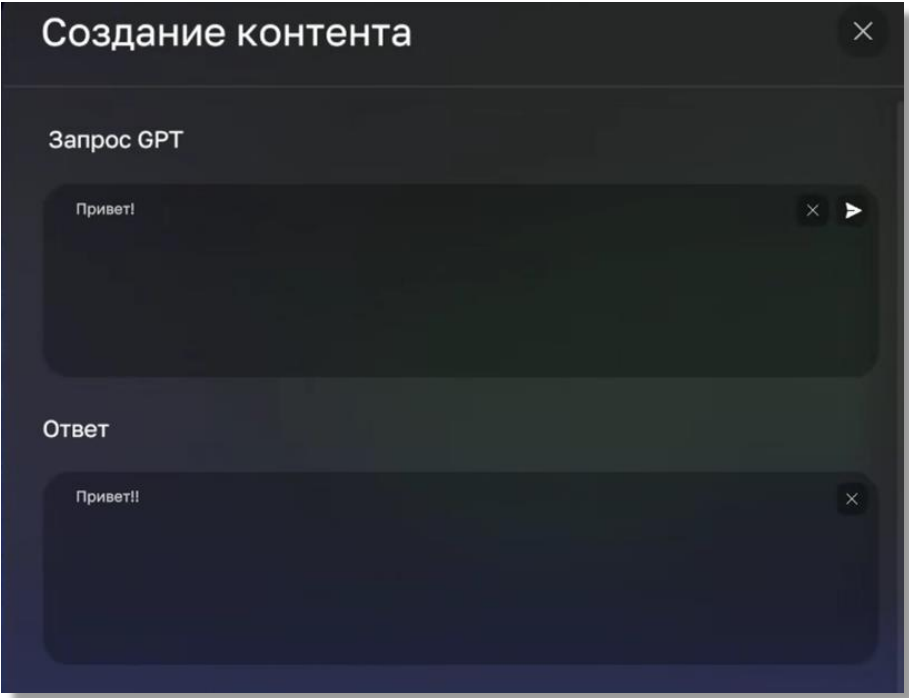


Рисунок 9 – Окно настройки запроса и ответа

- **Поле «Запрос GPT»** — используется для ввода вопроса, который может задать пользователь. Поддерживаются стандартные операции копирования и вставки: Ctrl + C, Ctrl + V, выделение текста — Ctrl + A.

Для генерации ответа нажмите кнопку запуска GPT (см. рисунок 10). Повторное нажатие формирует новые варианты ответов, позволяя выбрать наиболее подходящий.

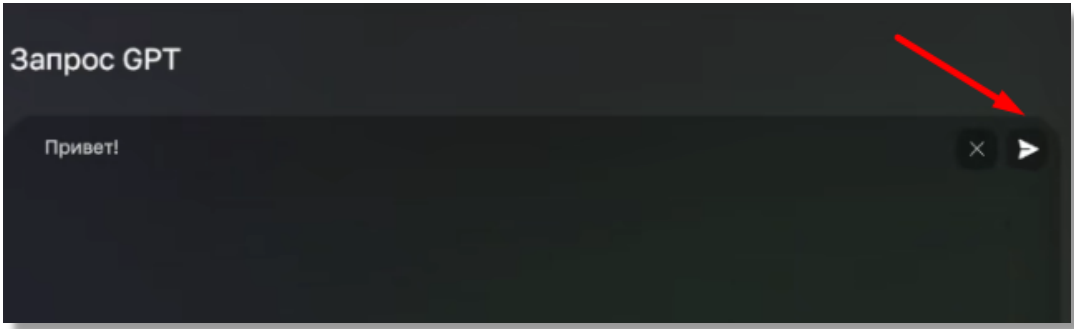


Рисунок 10 – Кнопка генерации ответа

Поле «Ответ» — содержит текст, автоматически сгенерированный моделью GPT. Также возможно ручное редактирование или замена текста (см. рисунок 11).

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

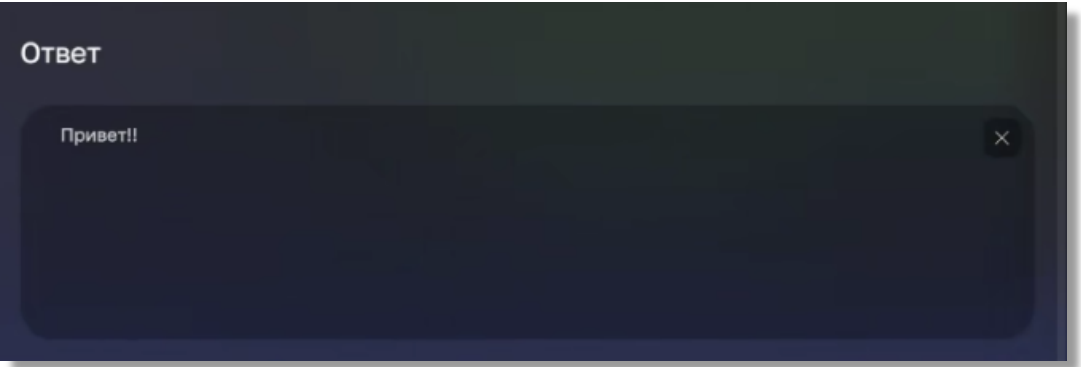


Рисунок 11 – Поле редактирования ответа

4.2.2 Настройка эмоций

В секциях «Эмоции в аудио» и «Эмоции в видео» (см. рисунок 12) указывается тип эмоциональной реакции ассистента при воспроизведении ответа.

Поддерживаются следующие типы эмоций:

- негативная,
- нейтральная,
- позитивная.

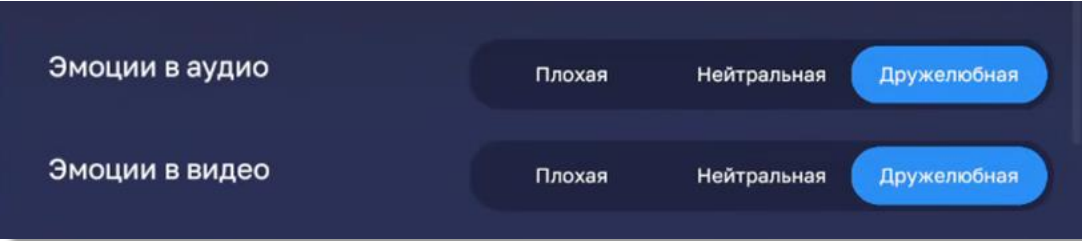


Рисунок 12 – Настройка эмоциональной реакции

4.2.3 Запись аудио- и видеоответов

Ассистент предоставляет возможность записи произносимого ответа в виде аудио- или видеофайла. Для этого:

- 1) Введите текст в поле «**Ответ**» или сгенерируйте его с помощью GPT;
- 2) Нажмите кнопку «**Аудио**» или «**Видео**» (см. рисунок 13);
- 3) Сгенерированные файлы будут автоматически сохранены в директорию:

C:\Users\aktay\OneDrive\Рабочий стол\ZVW\zvz_Data\AppData

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						23

C:\Users\aktay\OneDrive\Рабочий стол\ZVV\zvz_Data\AppData (рис. 14).

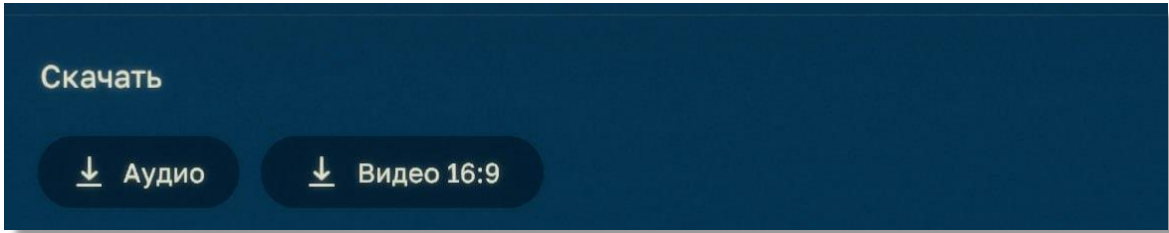


Рисунок 13 — Интерфейс скачивания файлов

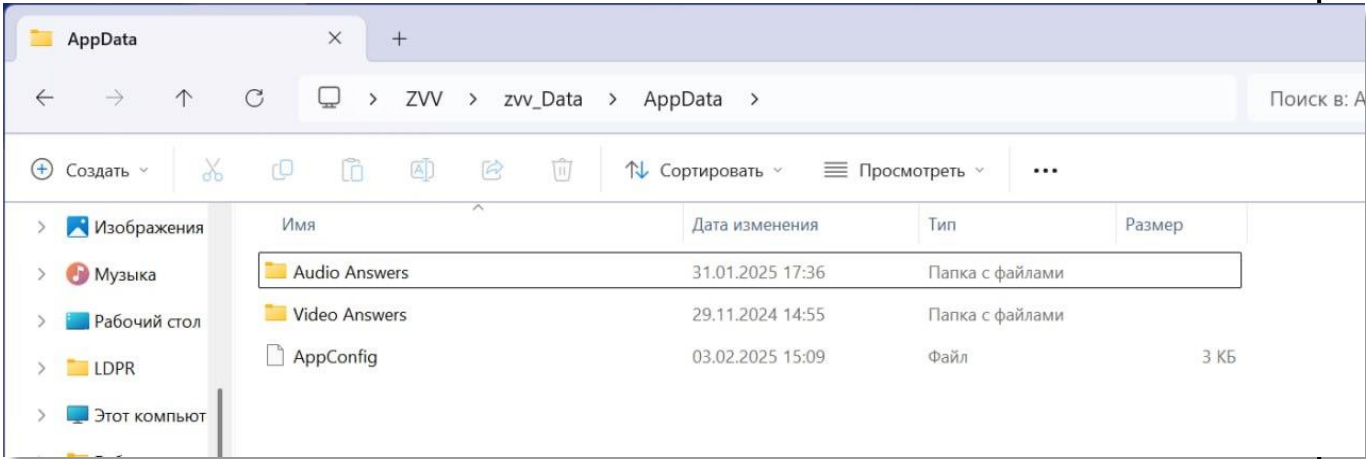


Рисунок 13 – Структура папок для сохранения аудио- и видеоответов

4.2.4 Настройка ассистента

Для доступа к настройкам нейросетевых модулей выполните следующие действия:

- 1) Перейдите в **режим администратора** (нажмите Ctrl + I);
- 2) правом верхнем углу выберите кнопку «**Все настройки**» (см. рисунок 8);
- 3) На экране отобразятся секции параметров (см. рисунок 14).

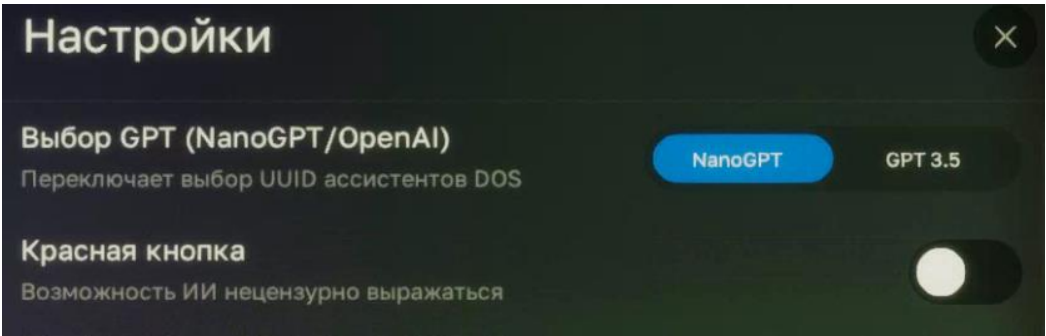


Рисунок 14 – Настройки нейросети

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

Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01	Лист
						24

В меню доступны следующие параметры:

- Выбор модели GPT – переключение между:
 - 1) внутренней моделью, разработанной ООО «Лаборатория Наносемантика» (NanoGPT);
 - 2) моделью от OpenAI.
- «Красная кнопка» – при активации разрешает использование нецензурной лексики в ответах ассистента.
- **Выбор имени пользователя** – позволяет задать обращение к пользователю по имени. Для этого введите соответствующий идентификатор (латиницей) в поле выбора персоны (см. рисунок 15).

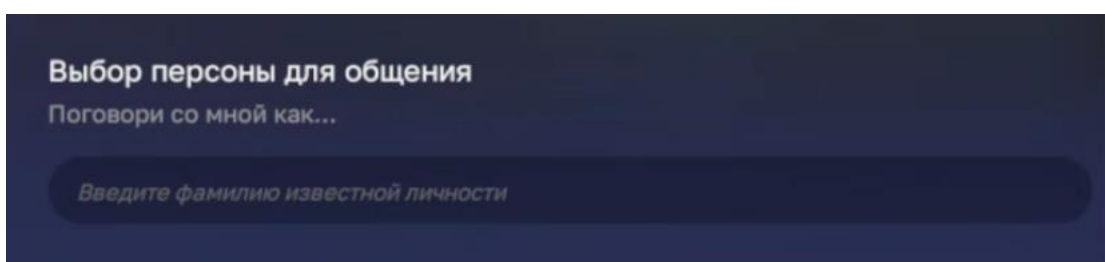


Рисунок 15 – Выбор персоны для общения

Поддерживаемые идентификаторы и соответствующие имена приведены в Таблице 5.

Таблица 5 – Список доступных персон

Идентификатор	Имя
gref	Герман Греф
kirienko	Сергей Кириенко
korchevnikov	Борис Корчевников
koshelev	Владимир Кошелев
mironov	Сергей Миронов
mishustin	Михаил Мишустин
peskov	Дмитрий Песков
popov	Евгений Попов
pozner	Владимир Познер
putin	Владимир Владимирович Путин
siluanov	Антон Силуанов
skabeeva	Ольга Скабеева
slutsky	Леонид Слуцкий

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата	<div>38264925.425000.676-0134 01</div>					Лист
										25
Изм.	Лист	№ докум.	Подп.	Дата						

sobchak	Ксения Собчак
solovyov	Владимир Соловьёв
volodin	Слава Володин
vorsobin	Владимир Ворсобин
zyuganov	Евгений Зюганов

Дополнительно в секции «Дополнительно» можно включить или отключить отображение субтитров для голосовых ответов ассистента (см. рисунок 16).

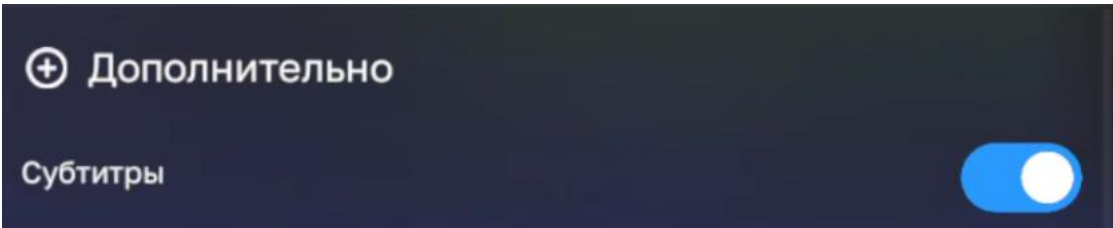


Рисунок 16 – Настройка субтитров

4.3 Процесс взаимодействия с ассистентом

После загрузки ассистент переходит в состояние ожидания: аватар отображает **анимации покоя (Idle)**.

При активации кнопки микрофона:

- ассистент переходит в активное состояние;
- аватар фокусируется в центр экрана;
- начинается воспроизведение **анимации размышления (Thinking)**.

Пользователь может говорить в микрофон. Запись речи отправляется в модуль **ASR**, при этом используется микрофон, выбранный по умолчанию в операционной системе.

Для определения наличия речи применяется модуль **VAD (Voice Activity Detection)**. Если в течение ~6 секунд после активации не зафиксирована голосовая активность, запись останавливается, и ассистент возвращается в режим покоя.

После окончания речи:

- микрофон отключается;
- проигрывается анимация «услышал вопрос» (**Heard**).

Если по завершении речи запись не остановлена (например, из-за фонового шума), её можно прервать вручную нажатием кнопки микрофона.

Далее процесс продолжается следующим образом:

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата	<div>38264925.425000.676-0134 01</div>					Лист
										26
Изм.	Лист	№ докум.	Подп.	Дата						

- 1) Текст, распознанный ASR, передаётся в **DialogOS**, где формируется ответ (с возможной генерацией через GPT);
- 2) Ответ и сопутствующие теги передаются в модуль **TTS**, где преобразуются в речь;
- 3) Голосовой ответ и субтитры (если активны) возвращаются в **Unity**;
- 4) Ассистент воспроизводит речь и **анимацию разговора (Talk)**;
- 5) По завершении ответа ассистент возвращается в **состояние покоя (Idle)**.

Прерывание и ошибки

- Повторное нажатие кнопки микрофона в любой момент (даже во время озвучивания ответа) приводит к прерыванию текущего действия и переходу к следующему вопросу.
- случае ошибки (например, отсутствия ответа от ASR или GPT) ассистент возвращается в состояние покоя без уведомлений.

4.4 Решение возможных проблем

4.4.1 Нарушение работы микрофона при его подключении/отключении

В случае подключения или отключения микрофона во время работы приложения ассистент может не распознать новое устройство. Возможные признаки:

- кнопка микрофона неактивна;
- кнопка активна, но индикатор громкости не отображается;
- голос не распознаётся.

Рекомендация: Перезапустите приложение и убедитесь, что в настройках операционной системы выбран корректный микрофон по умолчанию.

4.4.2 Отсутствие индикации громкости и распознавания речи

Если после активации микрофона не отображается индикатор уровня громкости и ассистент не реагирует на речь, выполните следующие шаги:

- 1) Проверьте, что выбранный микрофон установлен как основное устройство записи в операционной системе;

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата	<h3>4.4.1 Нарушение работы микрофона при его подключении/отключении</h3> <p>В случае подключения или отключения микрофона во время работы приложения ассистент может не распознать новое устройство. Возможные признаки:</p> <ul style="list-style-type: none">- кнопка микрофона неактивна;- кнопка активна, но индикатор громкости не отображается;- голос не распознаётся. <p>Рекомендация: Перезапустите приложение и убедитесь, что в настройках операционной системы выбран корректный микрофон по умолчанию.</p> <h3>4.4.2 Отсутствие индикации громкости и распознавания речи</h3> <p>Если после активации микрофона не отображается индикатор уровня громкости и ассистент не реагирует на речь, выполните следующие шаги:</p> <ol style="list-style-type: none">1) Проверьте, что выбранный микрофон установлен как основное устройство записи в операционной системе;					
Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01					Лист
										27

2) Убедитесь, что антивирус или система безопасности не блокируют доступ к микрофону.

Рекомендация: Добавьте приложение ассистента в список исключений или доверенных программ антивирусного ПО.

4.4.3 Микрофон активен, есть звуковая индикация, но речь не распознаётся

В случае, когда микрофон функционирует корректно (отображается пульсация и громкость), но система не распознаёт речь, возможны следующие причины:

- отсутствует подключение к сети Интернет;
- недоступны внешние модули ASR, DialogOS или TTS.

Рекомендация:

- 1) Убедитесь в наличии стабильного интернет-соединения;
- 2) Проверьте статус работы задействованных сервисов. При необходимости перезапустите соответствующие модули.

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата						
Изм.	Лист	№ докум.	Подп.	Дата	38264925.425000.676-0134 01					Лист
										28

СОКРАЩЕНИЯ, ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

Таблица 6 – Сокращения, термины и определения

Термин / Сокращение	Определение
Аватар	Анимированная 3D-модель, визуально имитирующая В. В. Жириновского. Используется для отображения мимики, эмоций и речевых анимаций ассистента.
Ассистент	Программный комплекс «ИИ Жириновский», реализующий голосовое и визуальное взаимодействие с пользователем, имитируя стиль речи В. В. Жириновского.
ИИ (Искусственный интеллект)	Совокупность алгоритмов, моделирующих интеллектуальное поведение: понимание, генерация, распознавание, обучение.
Косинусная близость	Мера сходства между двумя векторами, основанная на косинусе угла между ними. Применяется в модулях идентификации.
ASR (Automatic Speech Recognition)	Автоматическое распознавание речи. Преобразует голос пользователя в текст.
TTS (Text-to-Speech)	Синтез речи. Преобразует текст в голосовой аудиофайл.
GPT (Generative Pre-trained Transformer)	Языковая модель на базе нейросетей, использующая машинное обучение для генерации текстов.
DialogOS	Платформа для разработки и управления виртуальными ассистентами. Обеспечивает построение логики, подключение к каналам связи, аналитику и администрирование.
Unity	Игровой движок, используемый для отображения анимированного аватара и взаимодействия с пользователем.
VAD (Voice Activity Detection)	Технология определения начала и окончания речи в аудиопотоке.
Voice ID	Модуль голосовой идентификации. Сравнивает голос пользователя с сохранёнными шаблонами и определяет его личность.
Computer Vision	Модуль распознавания лиц. Использует алгоритмы компьютерного зрения для идентификации персон по изображениям или видеопотоку.
Dialog Flow	Сценарная логика взаимодействия с пользователем, построенная из последовательностей реплик, условий и реакций.
Docker	Платформа для упаковки, развертывания и изоляции компонентов ассистента в контейнерах.
Swagger UI	Графический интерфейс для тестирования REST API. Используется для ручной отладки запросов к модулям системы.

Инв. № подл.	Подп. и дата	Взам. инв. №	Инв. № дубл.	Подп. и дата
Изм.	Лист	№ докум.	Подп.	Дата
38264925.425000.676-0134 01				Лист
				29

Лист регистрации изменений

[illegible]

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

					<div> <div>38264925.425000.676-0134 01</div> <div>Лист</div> <div>30</div> </div>
Изм.	Лист	№ докум.	Подп.	Дата	