

D DIGIT / РЕЧЕВОЙ ИНТЕЛЛЕКТ / 

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

«Digit Речевой интеллект»

Руководство пользователя

Листов 13

СОДЕРЖАНИЕ

1	ВВЕДЕНИЕ.....	4
1.1	Общее описание системы.....	4
1.2	Переход в «Digit Речевой интеллект»	4
2	РАБОТА В ПО «DIGIT РЕЧЕВОЙ ИНТЕЛЛЕКТ»	5
2.1	Загрузка аудио и видеофайла	5
2.2	Запуск обработки файла	6
2.3	Просмотр результатов транскрибации.....	6
2.3.1.	Просмотр текста транскрипции	7
2.3.2.	Просмотр JSON-данных.....	8
3	АНАЛИЗ ТЕКСТА	10
3.1	Стандартный пайплайн.....	10
3.2	Анализ эффективности	11
3.3	Простое резюме.....	12

СПИСОК ТЕРМИНОВ И СОКРАЩЕНИЙ

Термин, сокращение	Значение
ПО	Программное обеспечение
ASR (Automatic Speech Recognition)	Технология автоматического распознавания речи, предназначенная для преобразования устной аудио- или видеозаписи в человеко-читаемый текст
Стенограмма (транскрипция)	Полный дословный письменный текст устной речи, полученный в результате преобразования аудио- или видеозаписи
Транскрибация	Автоматизированный процесс распознавания и записи устной речи из аудио- или видеоматериалов в текстовый формат
Резюмирование	Процедура создания структурированного краткого изложения основных итогов встречи или разговора на основе полной стенограммы
Суммаризация	Автоматическое формирование сжатой аннотации обсуждения (несколько предложений), передающей суть разговора в концентрированной форме
Протокол	Формализованный документ по заранее заданному шаблону, составляемый на основе транскрипта; включает полный или сокращённый текст разговора и, при необходимости, резюме или ключевые выводы

1 ВВЕДЕНИЕ

1.1 Общее описание системы

Программное обеспечение «Digit Речевой интеллект» представляет собой веб-приложение, реализованное на основе клиент-серверной архитектуры и предназначенное для автоматического преобразования аудио- и видеозаписей в текст с последующим интеллектуальным анализом содержания.

Программное обеспечение предназначено для выполнения следующих функций:

- загрузка аудио- и видеофайлов через веб-интерфейс;
- автоматическое распознавание речи;
- преобразование аудио- и видеозаписей в текстовую стенограмму;
- отображение результатов транскрибации в текстовом и JSON-формате;
- выполнение интеллектуального анализа текста;
- формирование аналитических отчетов по различным сценариям анализа.

1.2 Переход в «Digit Речевой интеллект»

Для начала работы необходимо открыть веб-приложение «Digit Речевой интеллект» в браузере по предоставленному URL-адресу.

2 РАБОТА В ПО «DIGIT РЕЧЕВОЙ ИНТЕЛЛЕКТ»

2.1 Загрузка аудио и видеофайла

Для загрузки файла пользователь может воспользоваться одним из двух способов:

- нажать кнопку «Browse files» и выбрать файл вручную;
- перетащить файл в область загрузки.

Область загрузки файла представлена на рисунке 1.

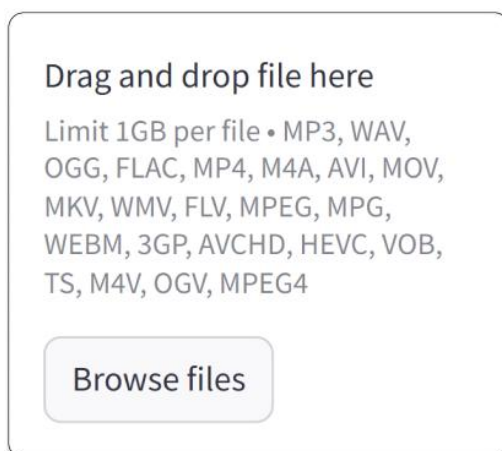


Рисунок 1 - Область загрузки файла

После выбора файла его имя отображается в интерфейсе системы, а кнопка «Начать обработку» становится активной (рисунок 2).

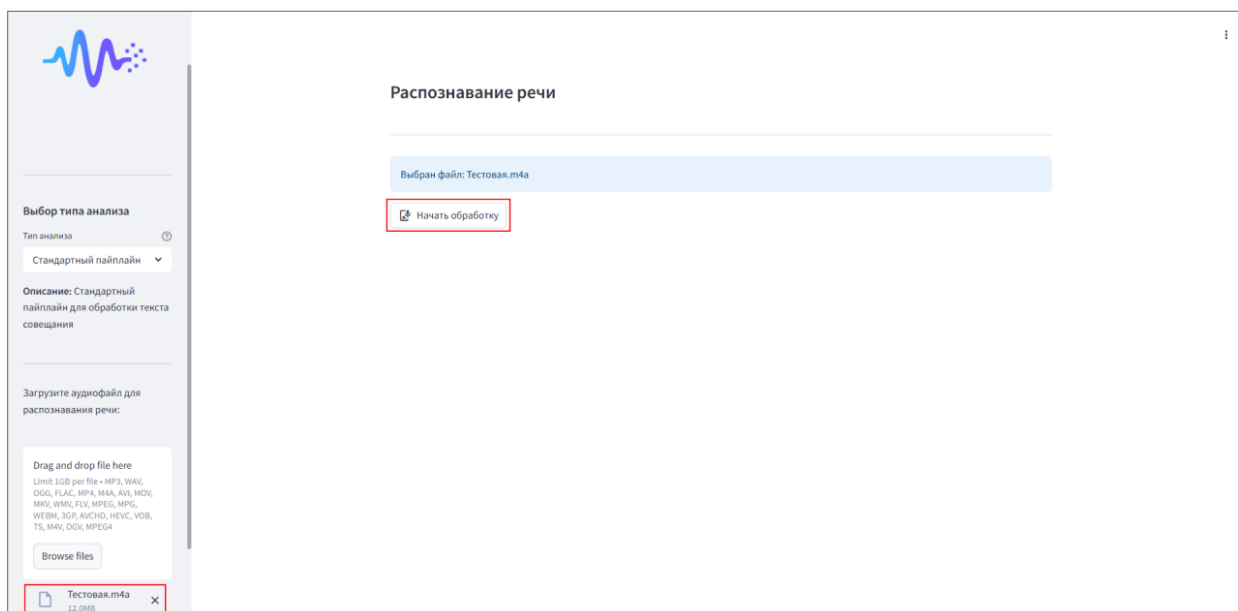


Рисунок 2 - Начать обработку загруженного файла

2.2 Запуск обработки файла

Для начала распознавания необходимо нажать кнопку «Начать обработку» (Рисунок 3).

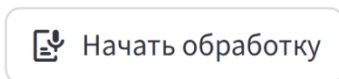


Рисунок 3 – Кнопка «Начать обработку»

После запуска система инициирует ASR-процесс, во время которого отображается окно статуса обработки (рисунок 4).

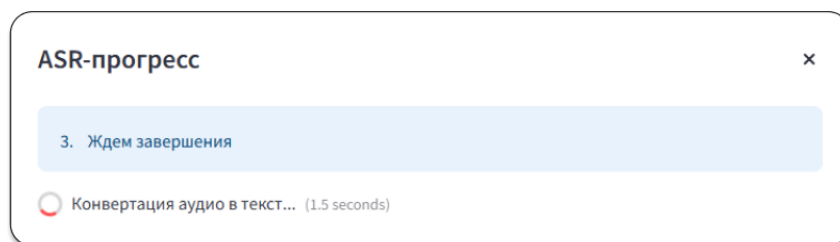


Рисунок 4 – Окно статуса обработки «Ждем завершения»

После успешного завершения процесса появляется сообщение: «Конвертация аудио в текст завершена» (рисунок 5).

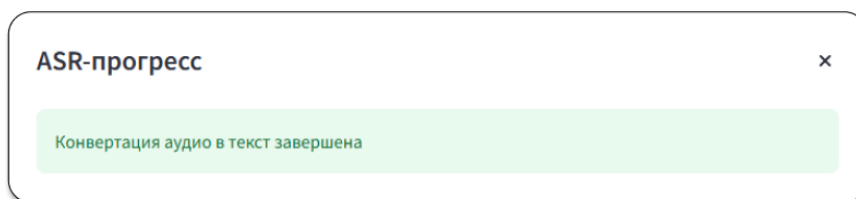
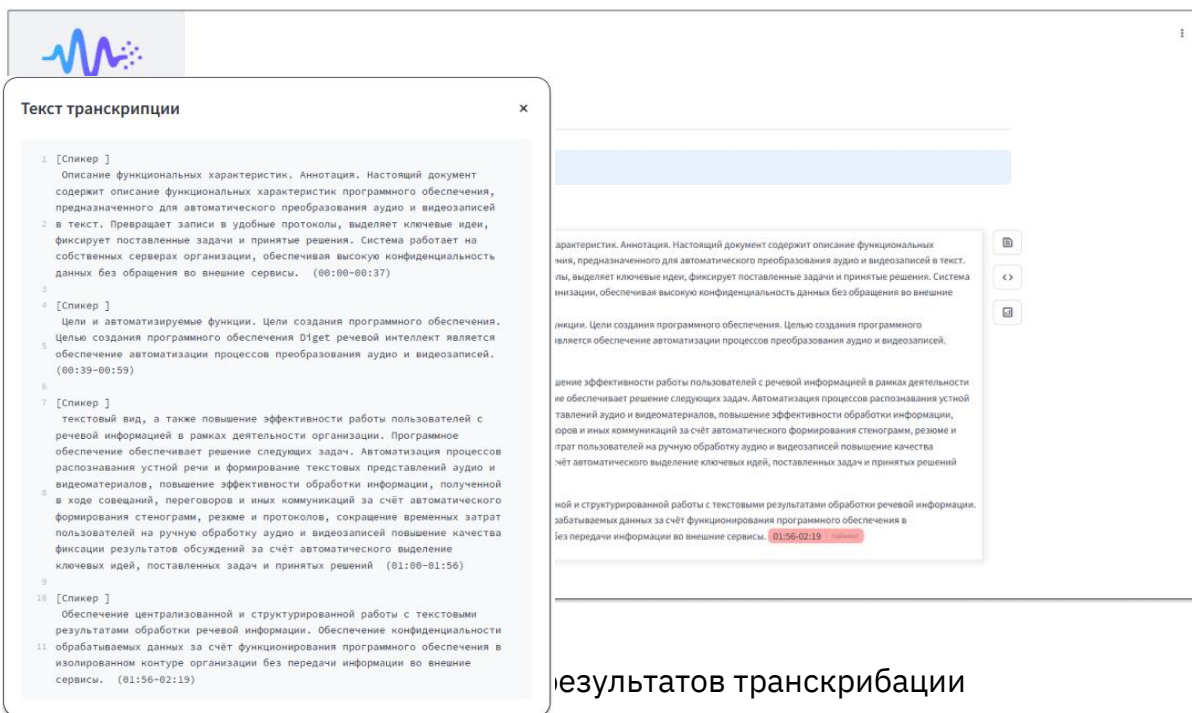


Рисунок 5 – Окно статуса обработки «Конвертация аудио в текст завершена»

2.3 Просмотр результатов транскрибации

После завершения распознавания пользователю выводится текст транскрибации (рисунок 6), а также доступны два дополнительных режима просмотра результатов:

- Вывод транскрибации в текстовом формате;
- Вывод транскрибации в формате JSON.



2.3.1. Просмотр текста транскрипции

Для просмотра стенограммы необходимо нажать кнопку **«Показать текст»** (рисунок

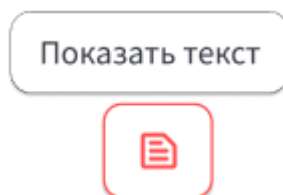


Рисунок 7 – Кнопка «Показать текст»

7).

После нажатия открывается окно, содержащее полный текст транскрипции аудиофайла (рисунок 8).

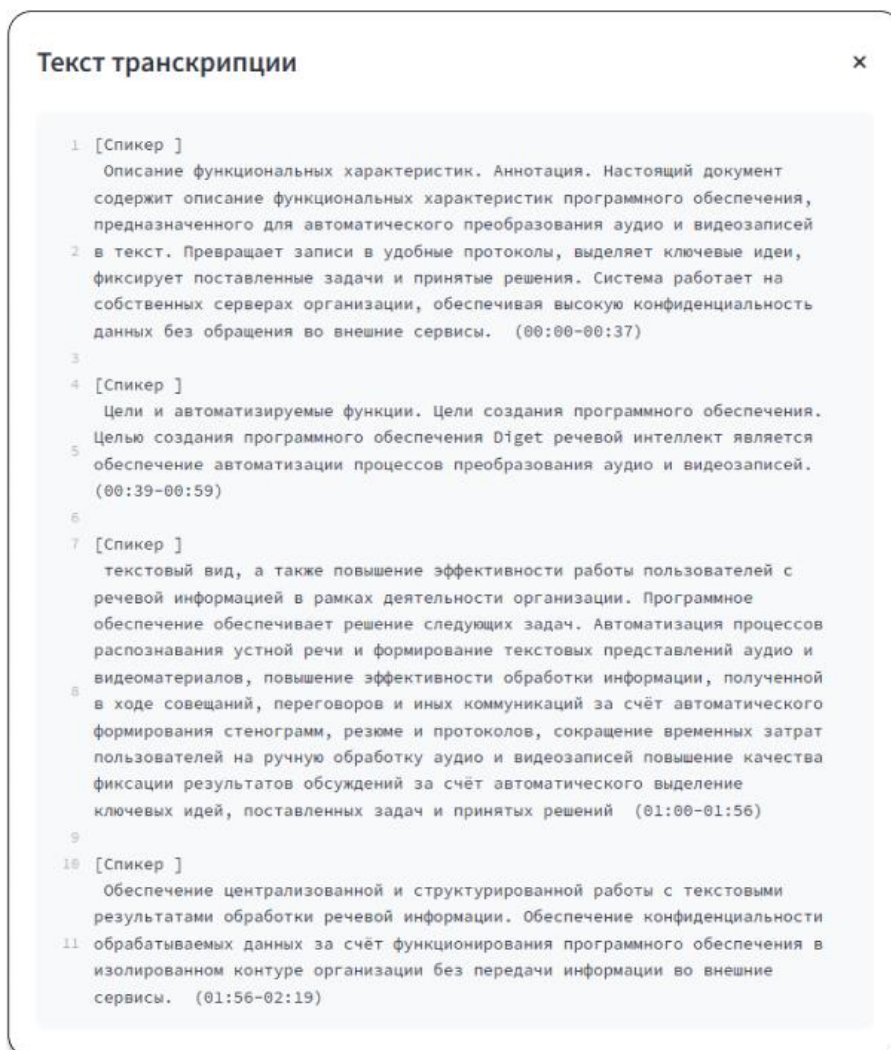


Рисунок 8 – Текст транскрипции

Пользователь может просматривать и копировать текст.

2.3.2. Просмотр JSON-данных

Для просмотра структурированных данных необходимо нажать кнопку «Показать JSON» (рисунок 9).

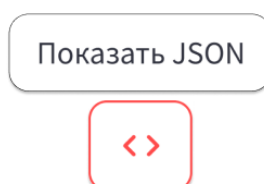


Рисунок 9 – Кнопка «Показать JSON»

После этого отображается JSON-представление результата обработки (рисунок 10).

JSON-данные содержат:

- распознанные текстовые сегменты;
- временные метки;

- вероятностные оценки.

```
JSON данные
{
  "result": [
    0: {
      "text":
        "Описание функциональных характеристик. Аннотация. Настоящий документ
        содержит описание функциональных характеристик программного
        обеспечения, предназначенного для автоматического преобразования аудио
        и видеозаписей в текст. Превращает записи в удобные протоколы, выделяет
        ключевые идеи, фиксирует поставленные задачи и принятые решения.
        Система работает на собственных серверах организации, обеспечивая
        высокую конфиденциальность данных без обращения во внешние сервисы."
      "start": 0.4
      "end": 37.1
    }
    1: {
      "text":
        "Цели и автоматизируемые функции. Цели создания программного
        обеспечения. Целью создания программного обеспечения Diget речевой
        интеллект является обеспечение автоматизации процессов преобразования
        аудио и видеозаписей."
      "start": 39.6
      "end": 59.7
    }
    2: {
      "text":
        "текстовый вид, а также повышение эффективности работы пользователей с
        речевой информацией в рамках деятельности организации. Программное
        обеспечение обеспечивает решение следующих задач. Автоматизация
        процессов распознавания устной речи и формирование текстовых
        представлений аудио и видеоматериалов, повышение эффективности
        обработки информации, полученной в ходе совещаний, переговоров и иных
        коммуникаций за счёт автоматического формирования стенограмм, резюме и
        протоколов, сокращение временных затрат пользователей на ручную
        обработку аудио и видеозаписей повышение качества фиксации результатов
```

Рисунок 10 – JSON-представление результата обработки

3 АНАЛИЗ ТЕКСТА

После завершения транскрибации пользователь может выполнить интеллектуальный анализ содержания записи.

Для этого необходимо:

- выбрать тип анализа в выпадающем списке;
- нажать кнопку «Анализ текста».

Выбор типа анализа показан на рисунке 11.

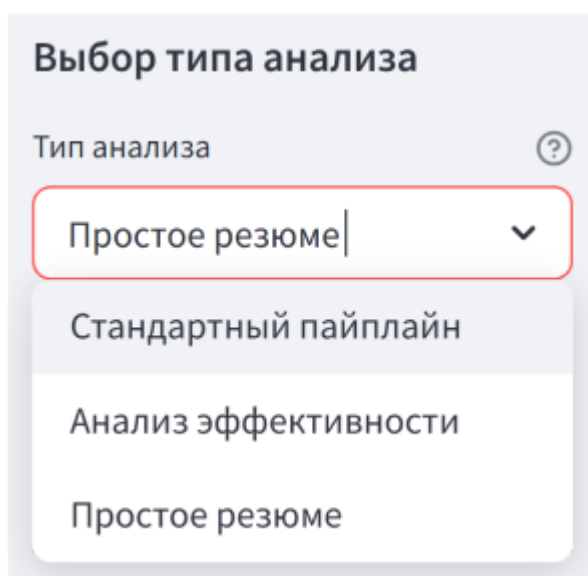


Рисунок 11 – Выбор типа анализа

Кнопка запуска анализа представлена на рисунке 12.

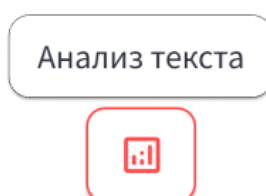


Рисунок 12 – Кнопка «Анализ текста»

3.1 Стандартный пайплайн

Режим «Стандартный пайплайн» предназначен для формирования структурированного аналитического отчета по содержанию совещания.

Система автоматически выделяет:

- основную цель совещания;
- второстепенные цели;

- тематические этапы;
- содержание обсуждения;
- вопросы и ответы;
- выводы и соглашения.

Пример результата анализа приведен на рисунке 13.

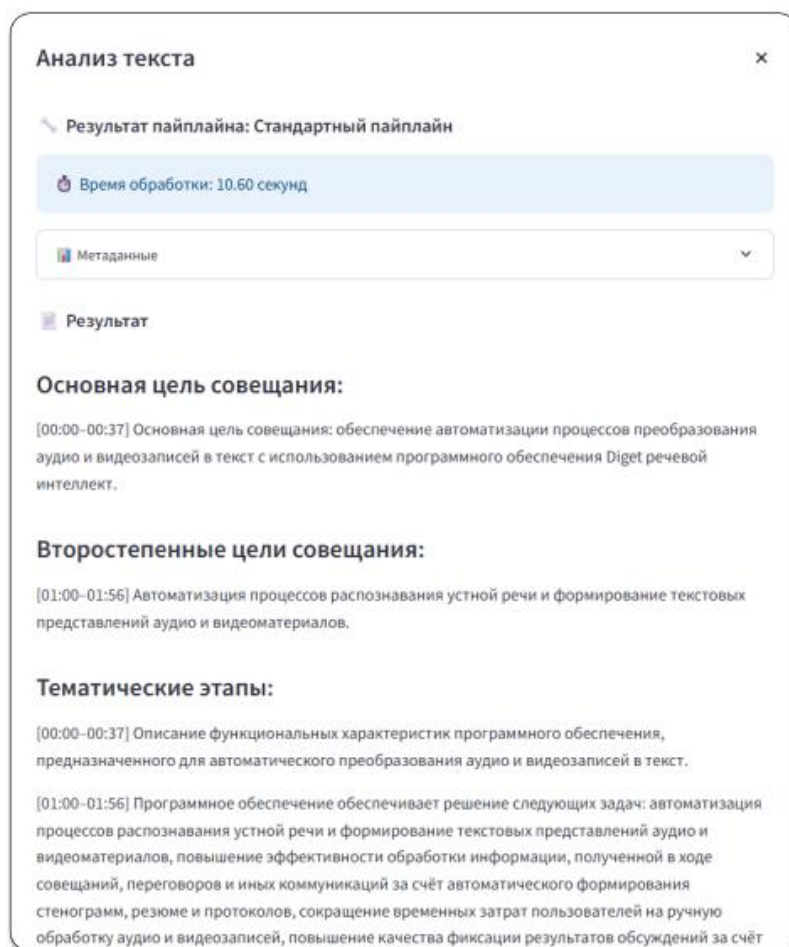


Рисунок 13 – Вывод результатов анализа «Стандартный пайплайн»

3.2 Анализ эффективности

Режим «Анализ эффективности» предназначен для оценки продуктивности встречи.

Система отображает следующие показатели:

- общая продолжительность совещания;
- количество принятых решений;
- количество заданных вопросов;
- показатель продуктивности в процентах.

Также отображается гистограмма распределения времени по категориям:

- разговор не по теме;

- административные вопросы;
- обсуждение;
- повестка;
- решения.

Пример результата анализа приведен на рисунке 14.

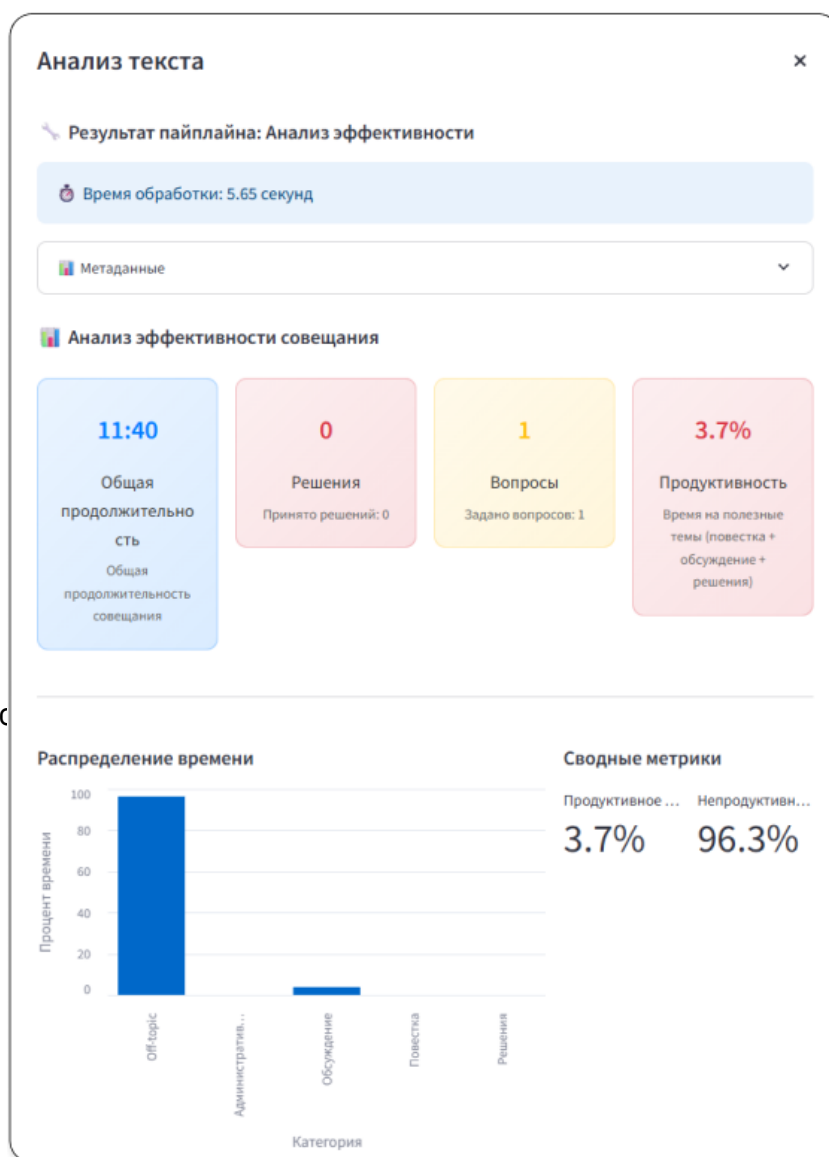


Рисунок 14 – Вывод результатов анализа «Анализ эффективности»

3.3 Простое резюме

Режим «Простое резюме» формирует краткое текстовое описание итогов совещания.

Результат включает:

- краткое резюме встречи;
- ключевые моменты;

- временные блоки обсуждения.

Пример результата анализа приведен на рисунке 15.

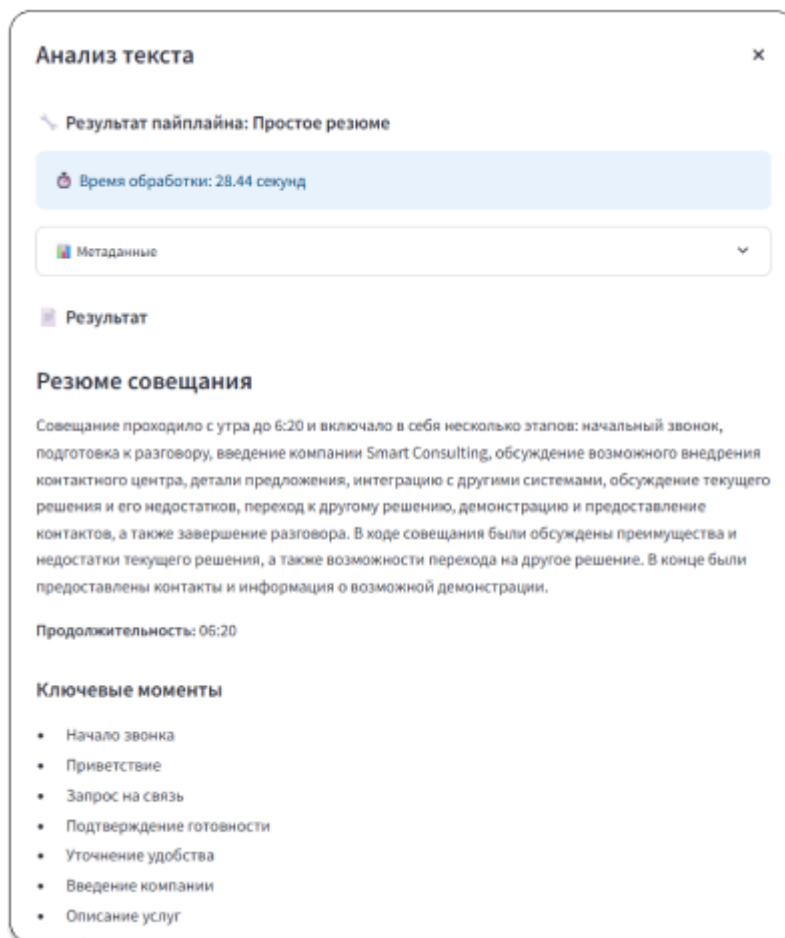


Рисунок 15 – Вывод результатов анализа «Просмотр резюме»