

MOSCOW FORENSICS DAY '25

12.09.25 г.

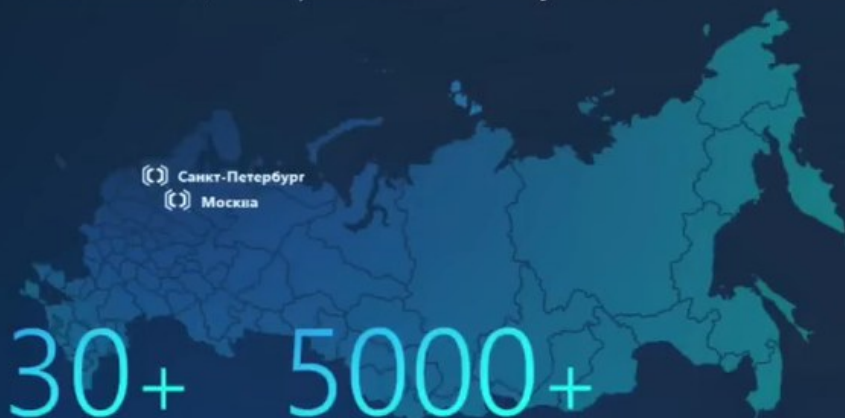
От аудио- и видеоданных к доказательствам

AI-аналитика, биометрия, детекция спуфинга и дипфейков

Спикер: Михаил Инкин
Руководитель группы проектов

О компании

Группа компаний ЦРТ — разработчик продуктов и решений на основе интеллектуальных речевых технологий, генеративного искусственного интеллекта, лицевой и голосовой биометрии



30+

лет на рынке технологий ИИ

5000+

проектов по всему миру, в т.ч. проекты национального масштаба



Один из крупнейших игроков рынка ИИ в России



Победитель международных научных конкурсов по распознаванию речи и биометрии

NIST

CHiM
CHALLENGE

ASVspoof



750+

сотрудников в штате компании

70+

команд R&D

20

кандидатов и докторов наук

Кадровый ресурс

ИТМО



Корпоративная лаборатория
Технологии человеко-машинного
взаимодействия

Решения LLM от Сбера

Трансформация бизнес-процессов



Поиск по базе
знаний



Расчёт
маржинального
дохода



Прогнозная
диагностика
оборудования



Ассистент в
закупках



GIGA IDE



GIGA CODE



GIT VERSE

Постоянно растущие объемы неструктурированных данных



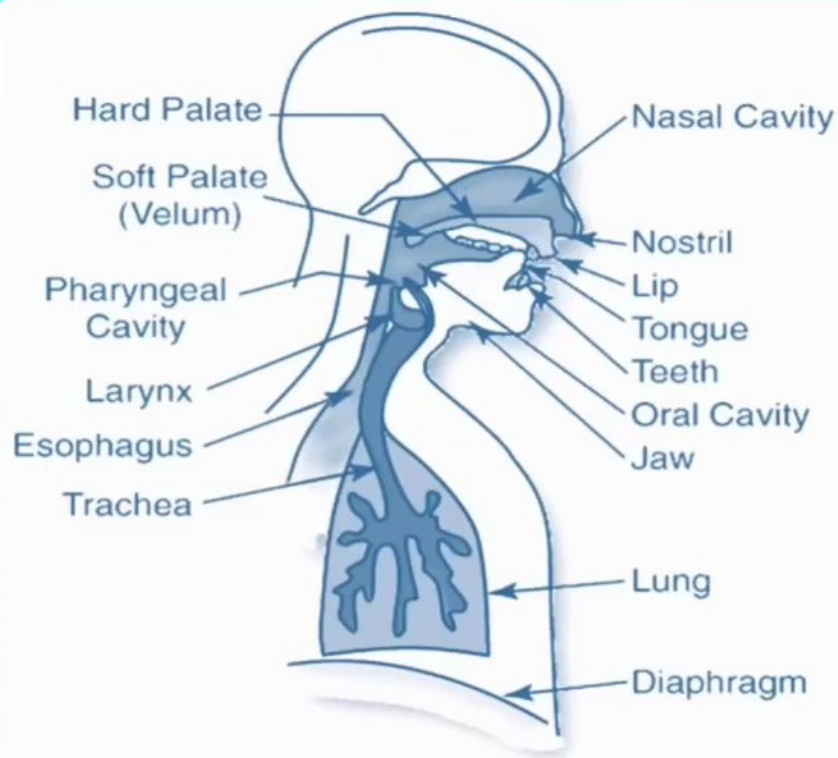
- ❗ Низкая эффективность поиска и обработки данных
- ❗ Высокая стоимость масштабирования систем обработки данных
- ❗ Низкая скорость анализа и принятия решений
- ❗ Потеря важной информации и ценных сведений

Типы биометрических систем



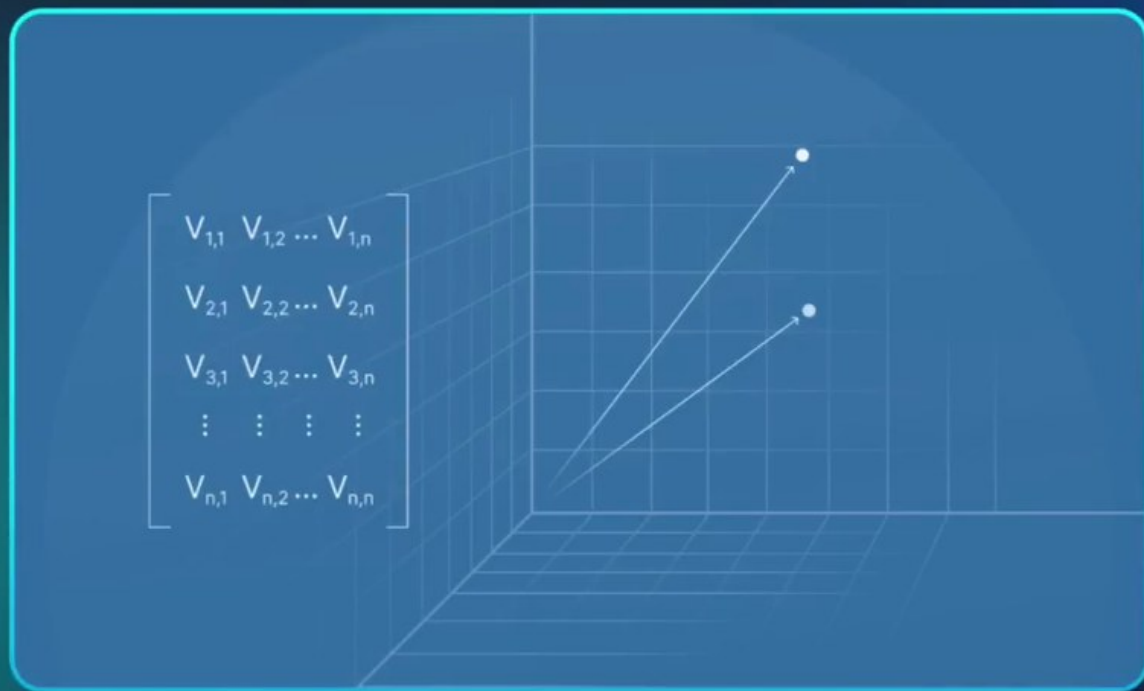
Синтез речи человеком

- Голос каждого человека уникален!
- Более 70 частей тела вовлечены в формирование голоса
- Важно не то, что мы говорим, а кто говорит
- Более 30 лет опыта использования для судебной экспертизы

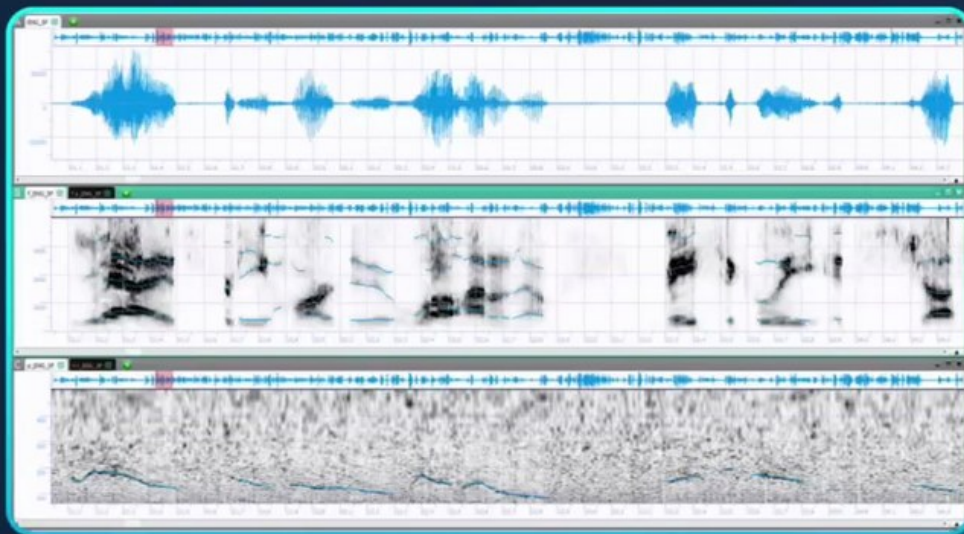


Модель диктора

Вектор признаков, эмбединг



Идентифицирующие характеристики голоса



Частота основного тона (F0) (ЧОТ, фундаментальная частота, основной тон) — частота колебаний голосовых складок определяется характеристиками речевого тракта и лежит в диапазонах (60-180 Гц М, 160-350 Гц Ж, и 200-650 Гц для детей).

Форманты — частотные максимумы, возникающие в результате резонанса речевого тракта человека.

Обычно три форманты (F1, F2, F3) достаточны как для звуко различения, так и для узнавания диктора.

Каждая форманта является функцией всех частей речевого тракта, хотя степень их влияния неодинакова в каждом случае.

Сложности задачи

Причины вариативности данных¹



Дикторские



Технологические

Чтение,
спонтанная
речь,
монолог,
диалог,
групповое
общение
...

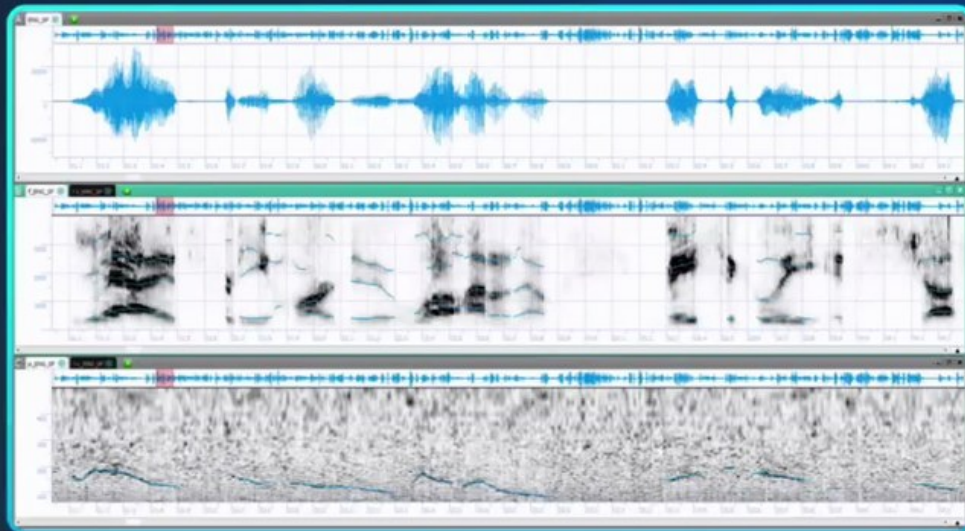
Коммуникативные



[1] - Speaker recognition by machines and humans [Hansen J.H.L., Hasan T., 2015]

Подходы к сравнению дикторов

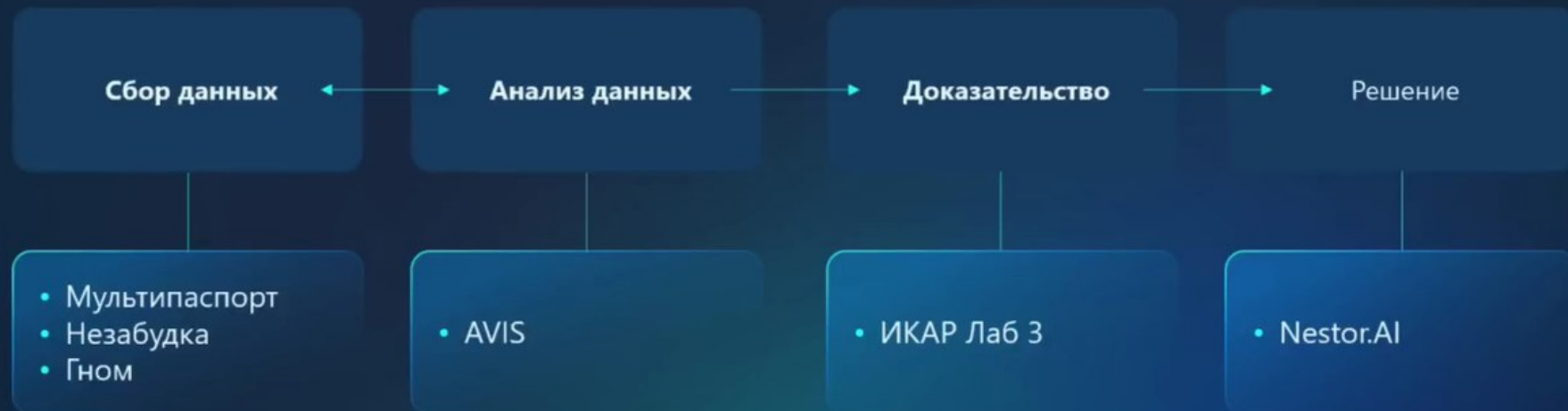
- Аудитивный
- Акустико-фонетический
- Спектрографический:
сравнение формант, ОТ, гармоник ОТ
- Автоматический



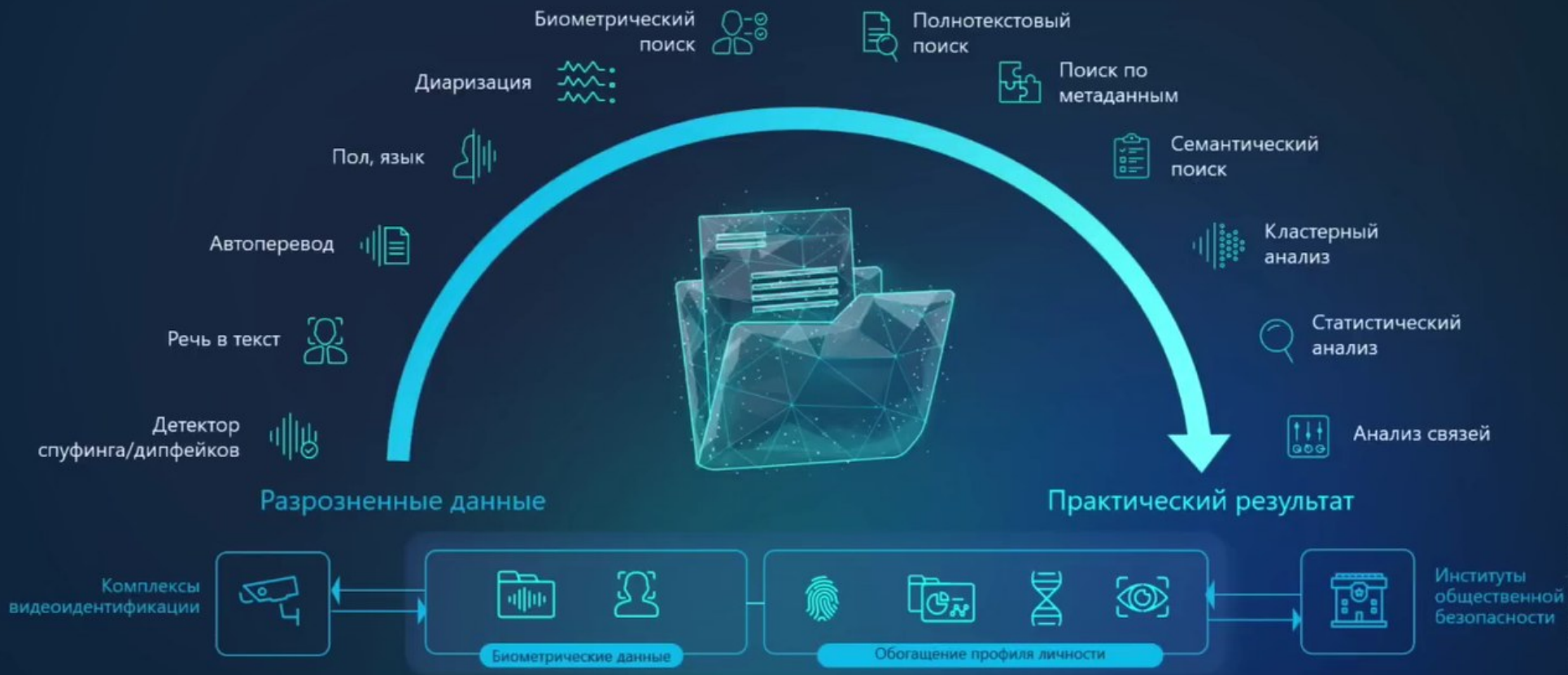
Решение прикладного кейса



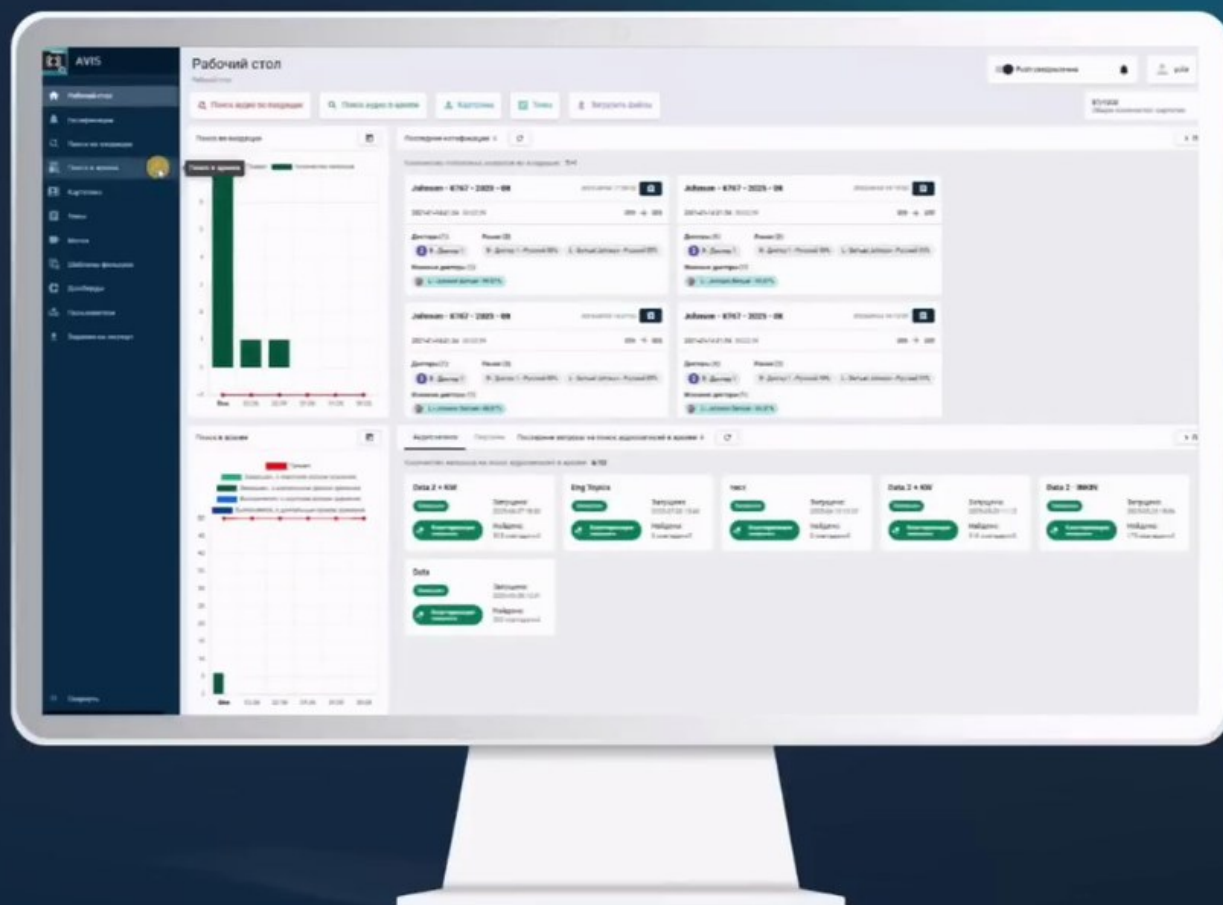
Продуктовая линейка и этапы криминалистического процесса



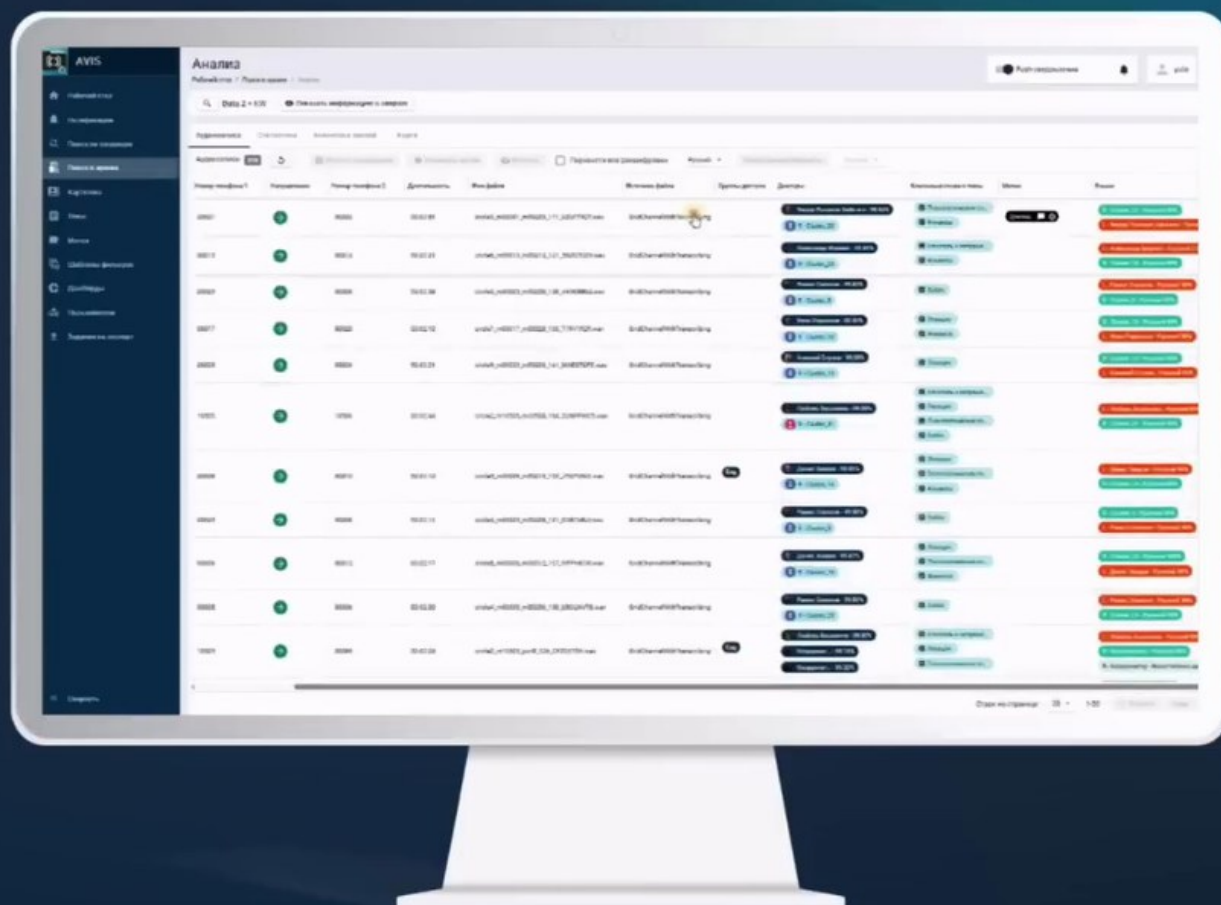
Продукты для общественной безопасности



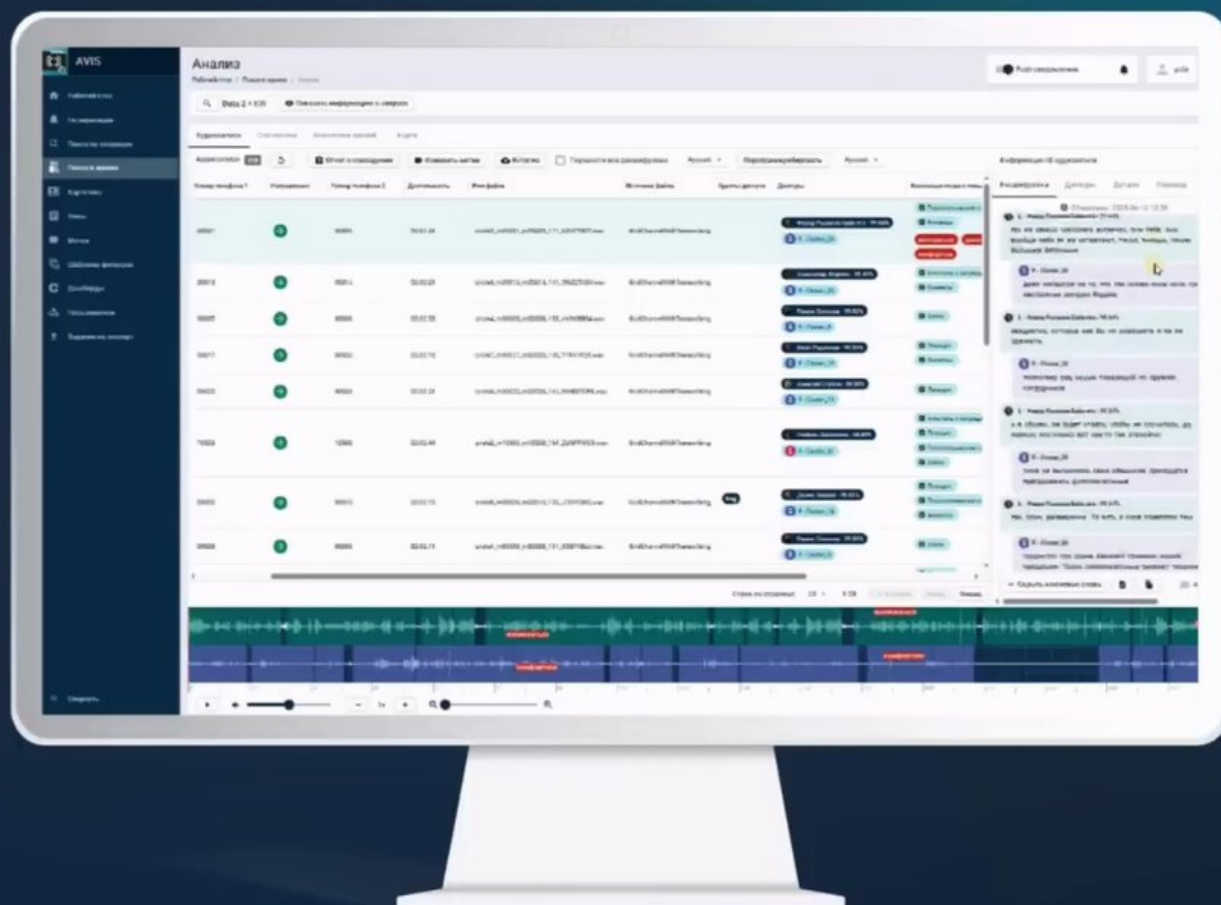
Что происходит с данными в AVIS



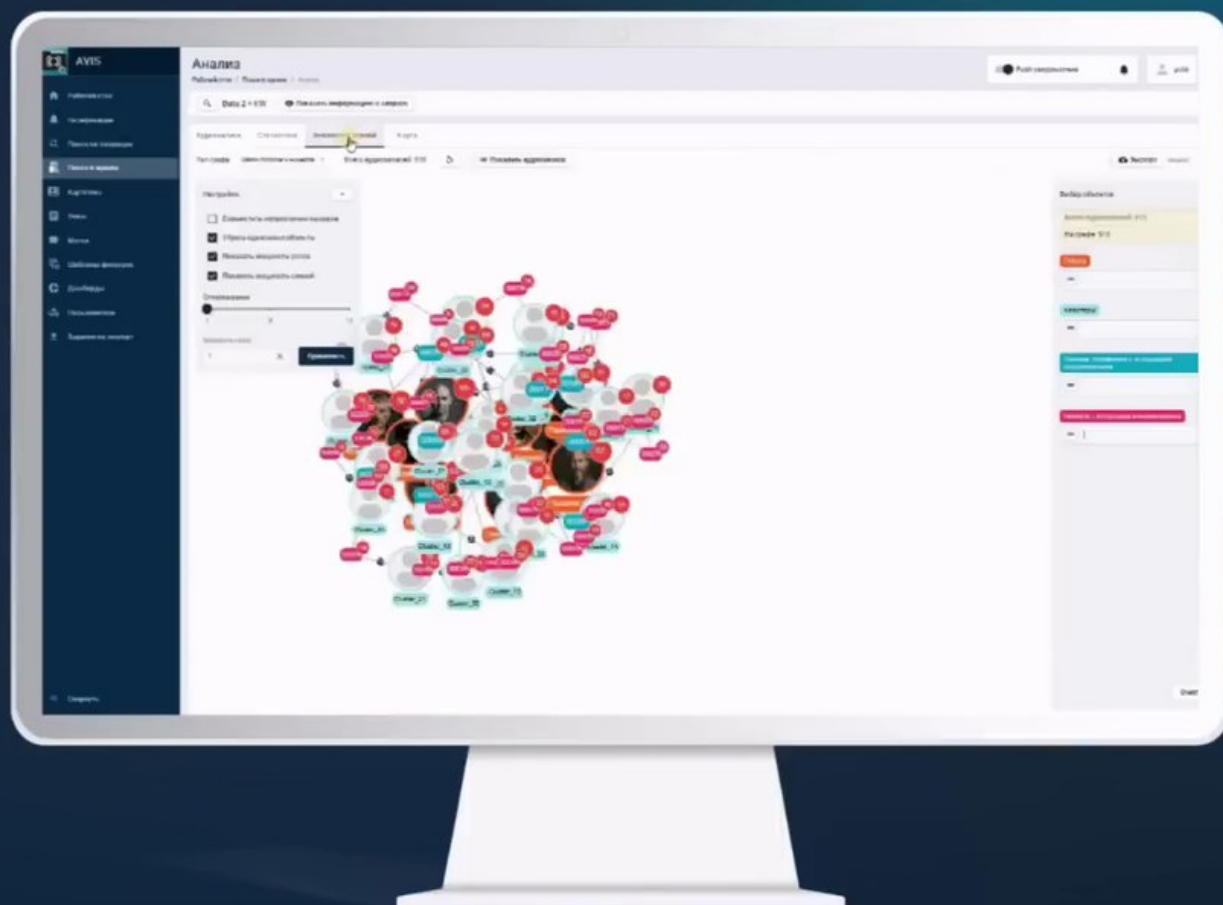
Что происходит с данными в AVIS



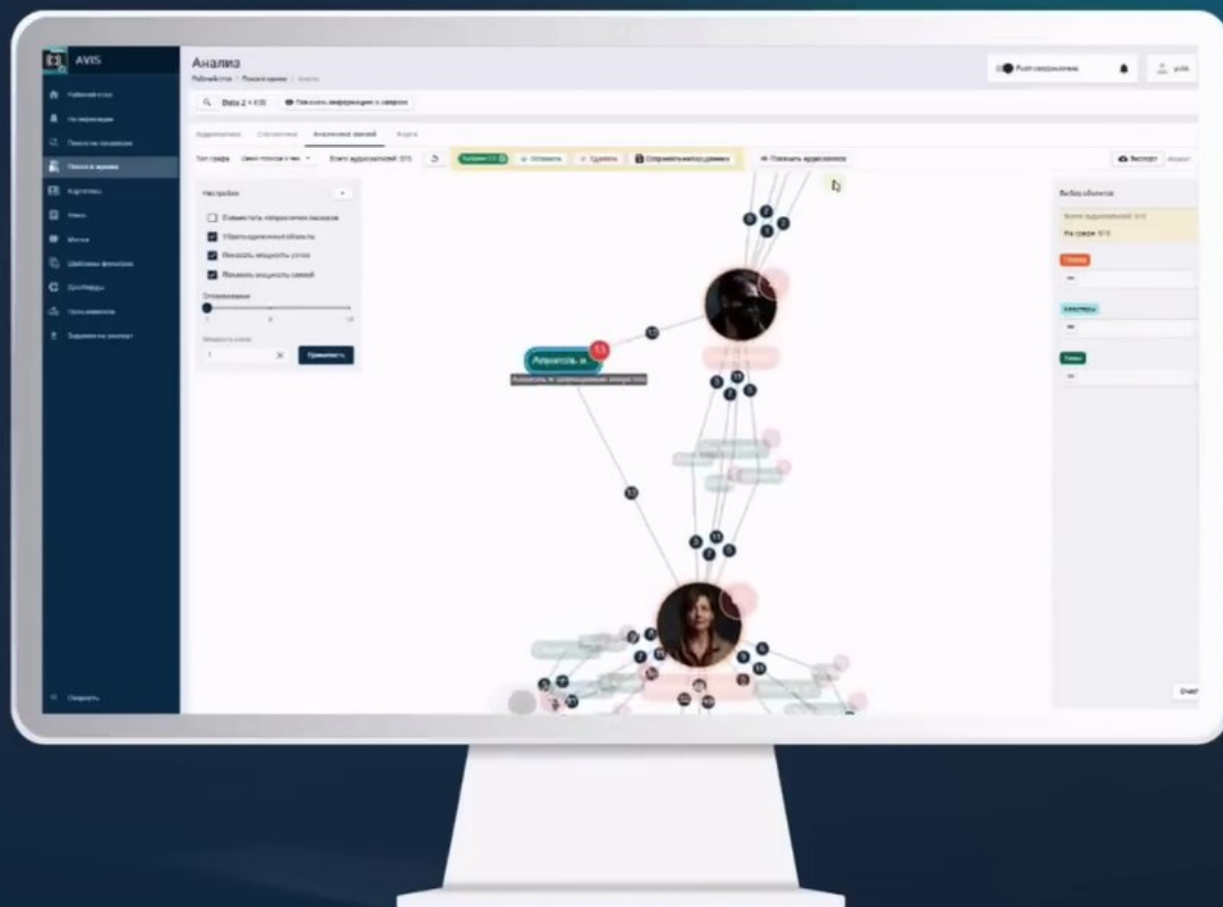
Что происходит с данными в AVIS



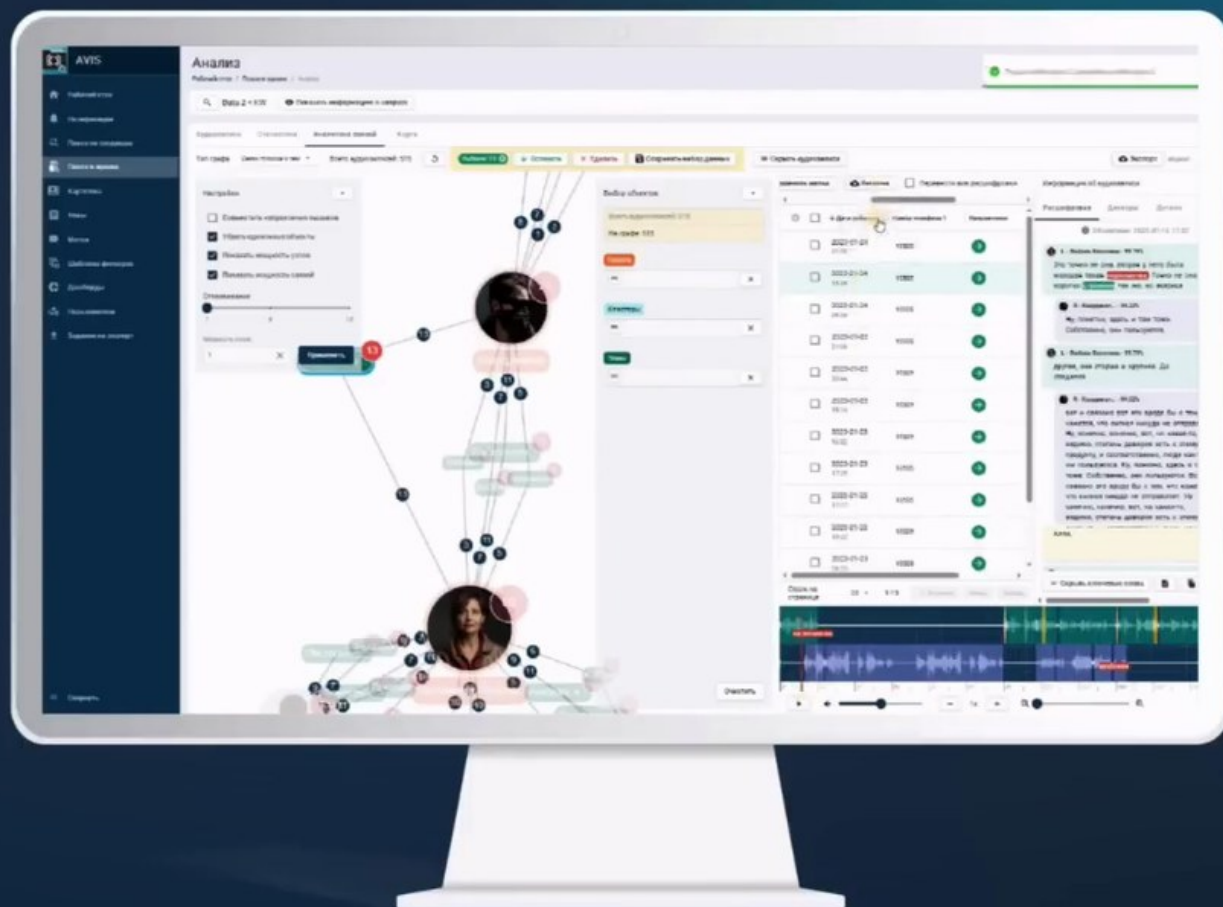
Что происходит с данными в AVIS



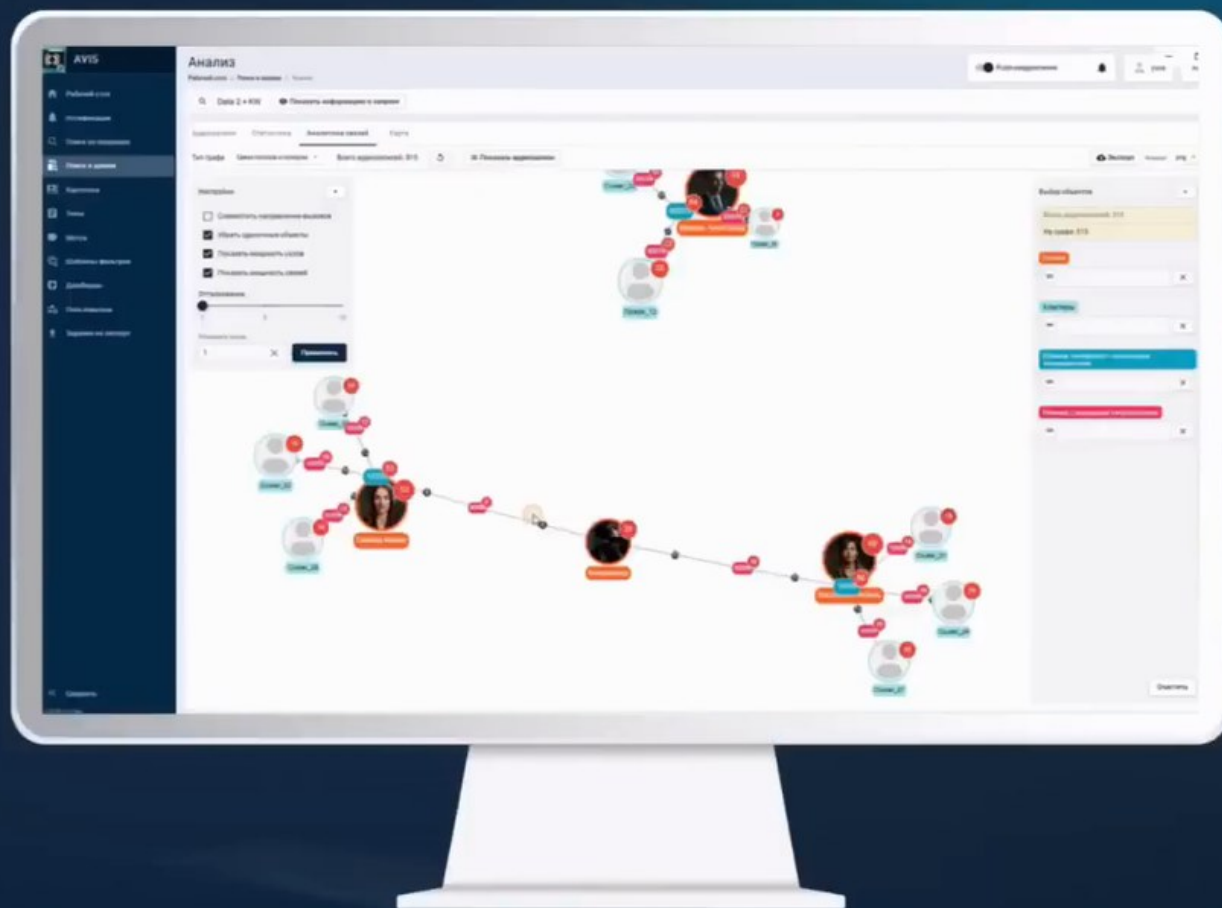
Что происходит с данными в AVIS



Что происходит с данными в AVIS



Биометрический граф связей



LLM в решении задач криминалиста

Генерация текста по инструкции

Создание связанного и грамматически правильного текста на заданную тему

Ответы на вопросы

Предоставление ответов на вопросы, основанные на известной LLM информации и контексте

Классификация текста

Определение категорий или меток для заданного текста (например, анализ тональности)

Персонализация ответов

Адаптация ответов и поведения под конкретного пользователя

Извлечение информации

Выделение ключевых фактов и данных из текста

Поддержка беседы

Ведение диалогов с пользователями, понимание и контекстуализация предыдущих реплик

Генерация и отладка кода

Написание и исправление ошибок в коде на различных языках программирования на основе описания задачи

Резюмирование текста

Краткое изложение длинных текстов, выделение основных идей

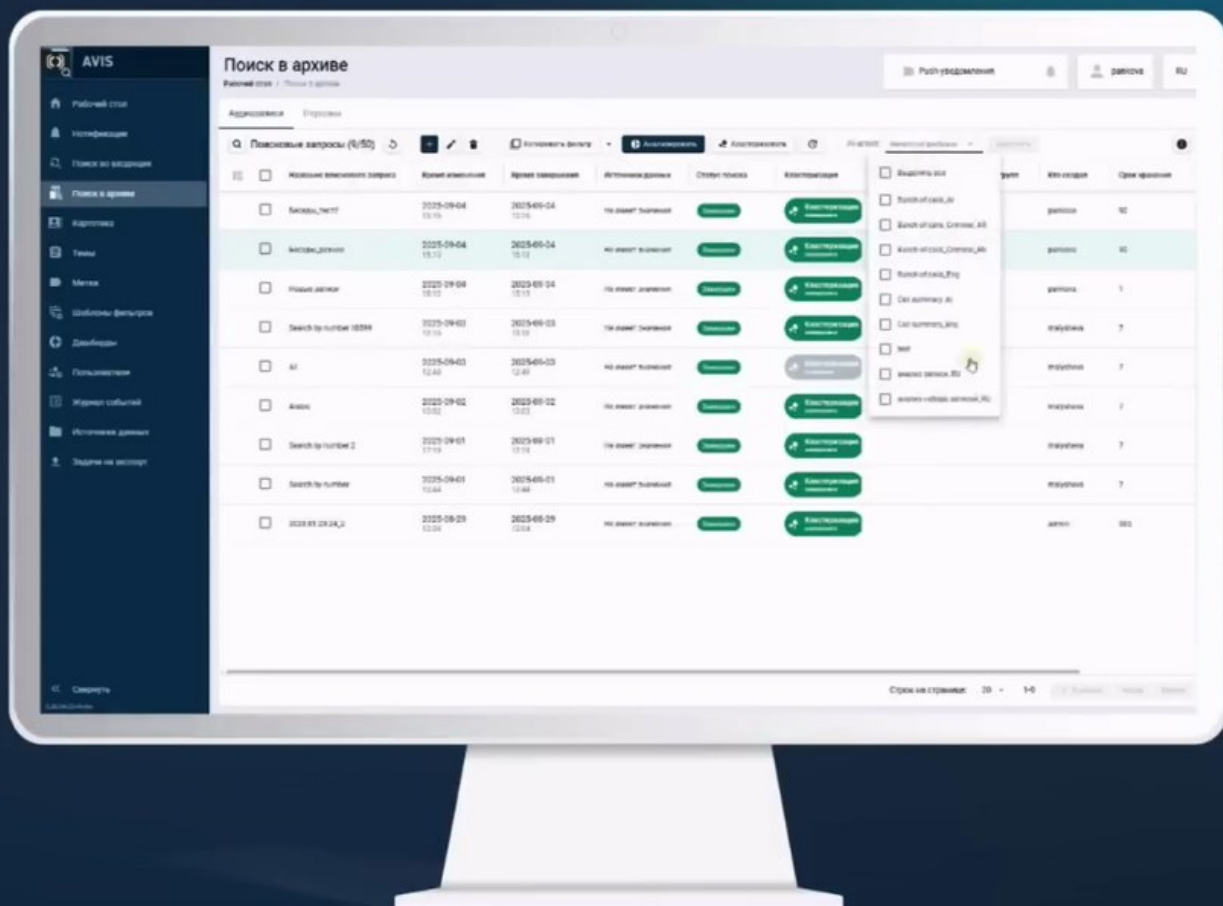
Перевод текста

Перевод текста с одного языка на другой

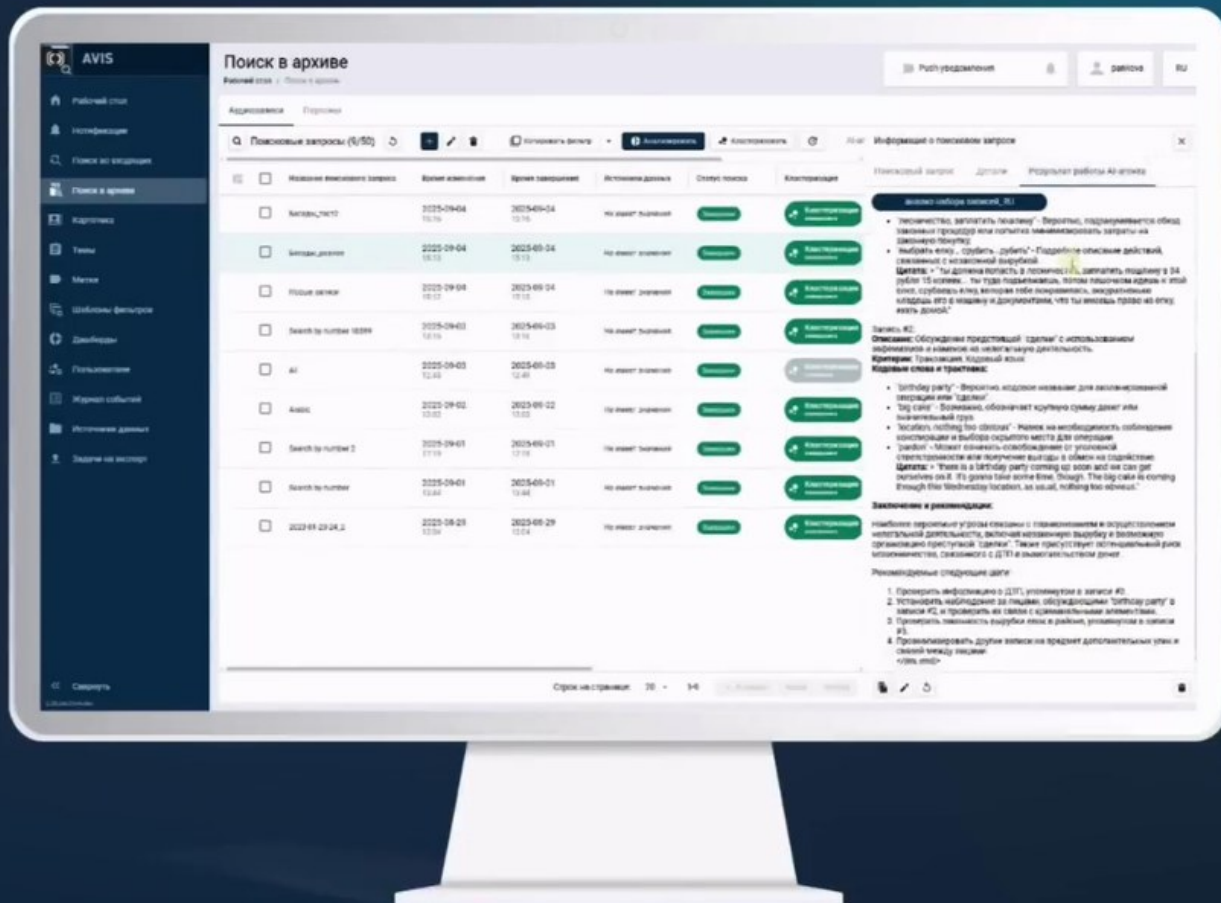
Интерпретация данных

Объяснение результатов анализа данных, визуализация данных

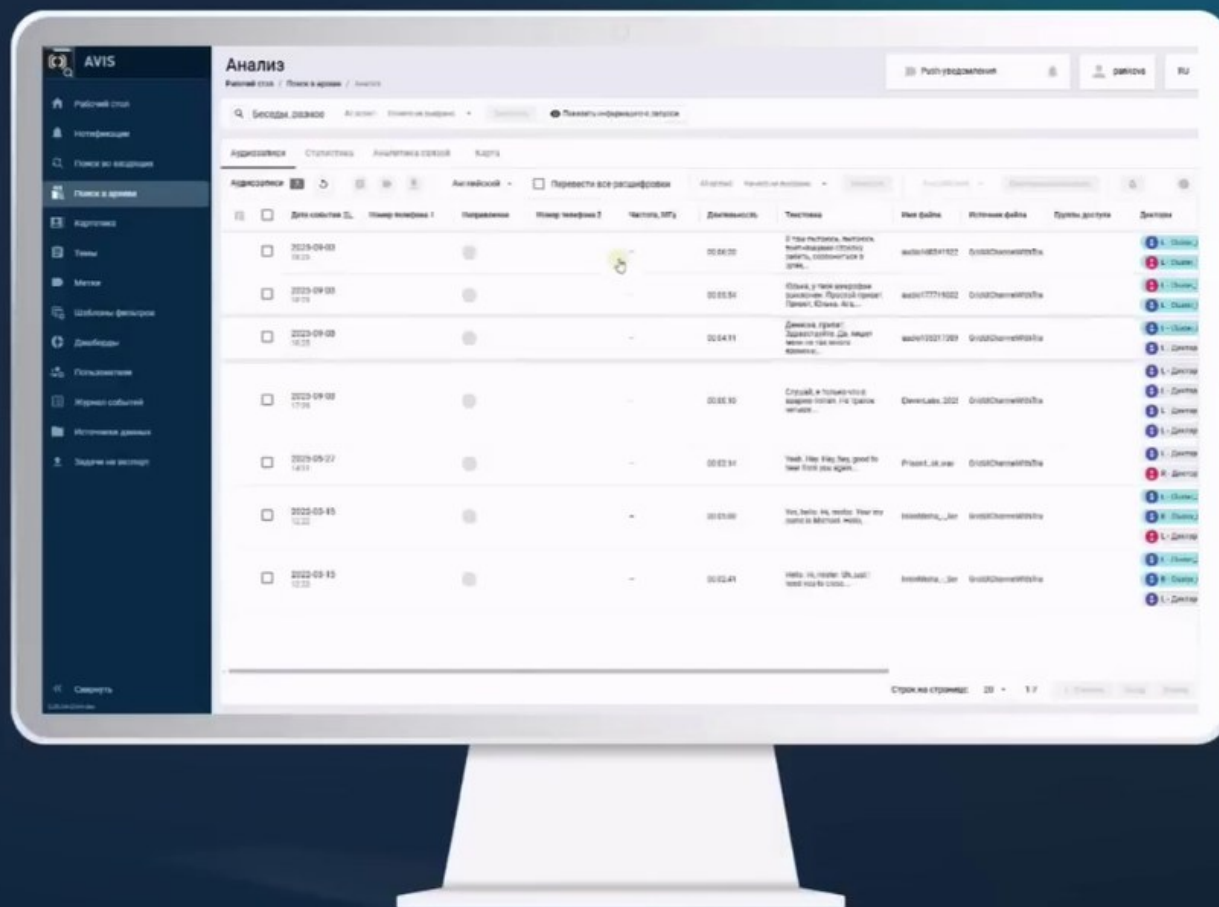
Массовая обработка информации



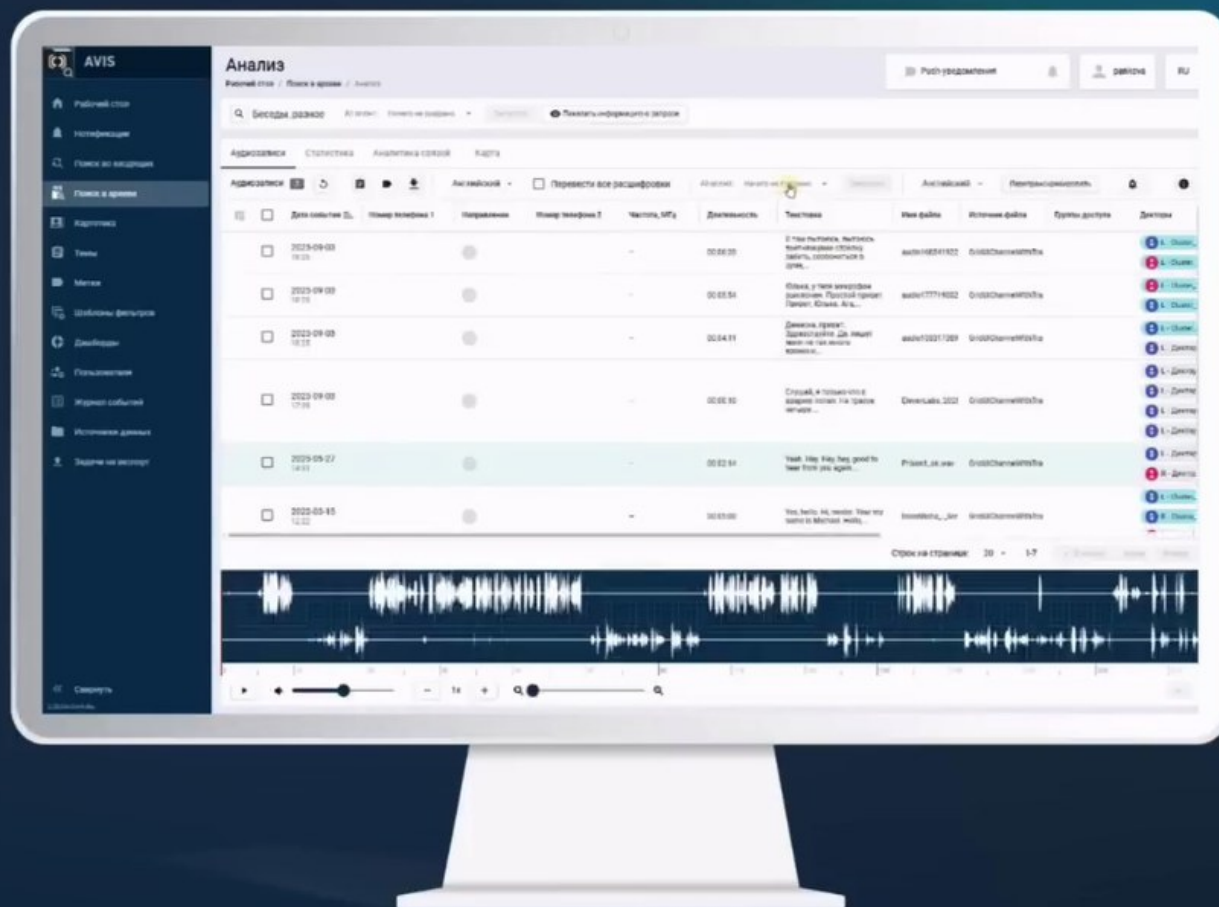
Массовая обработка информации



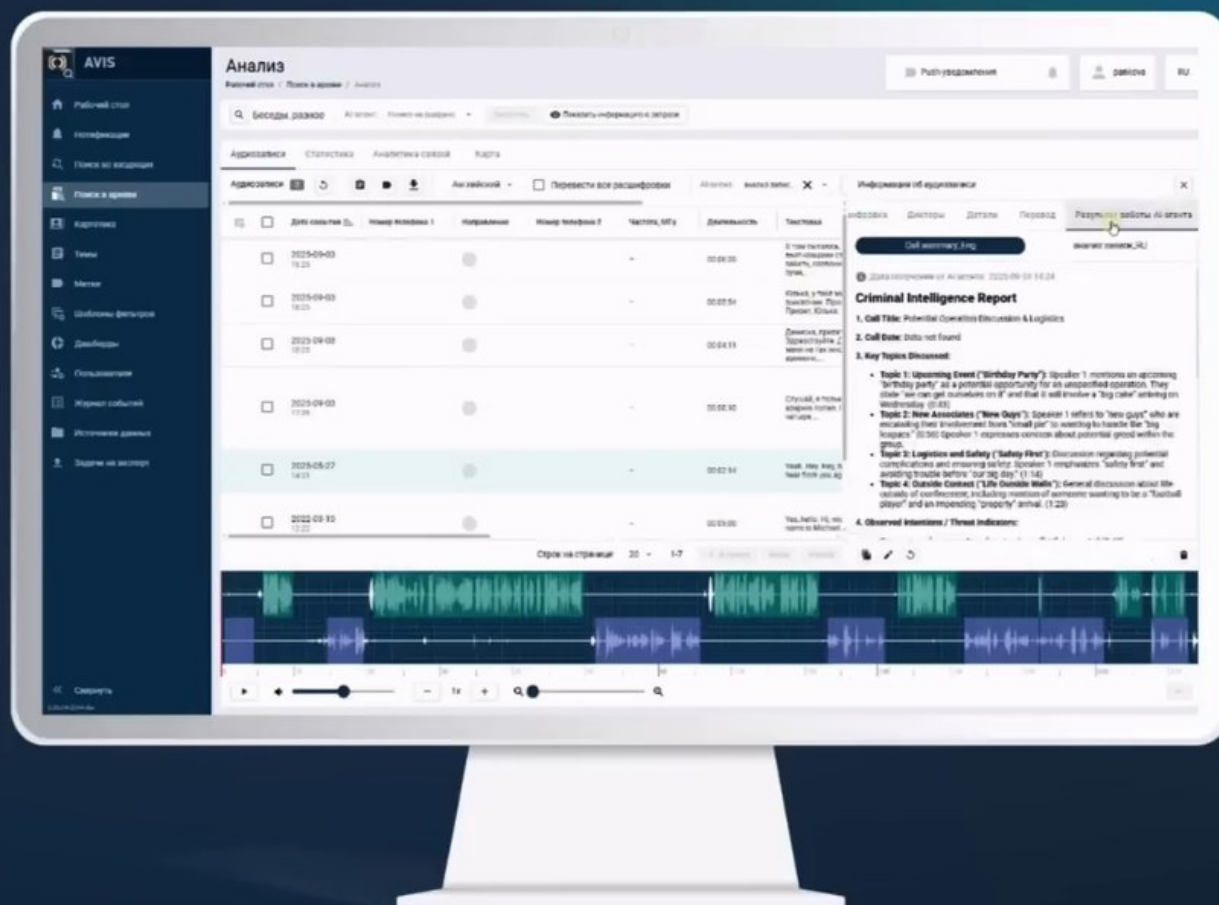
Массовая обработка информации



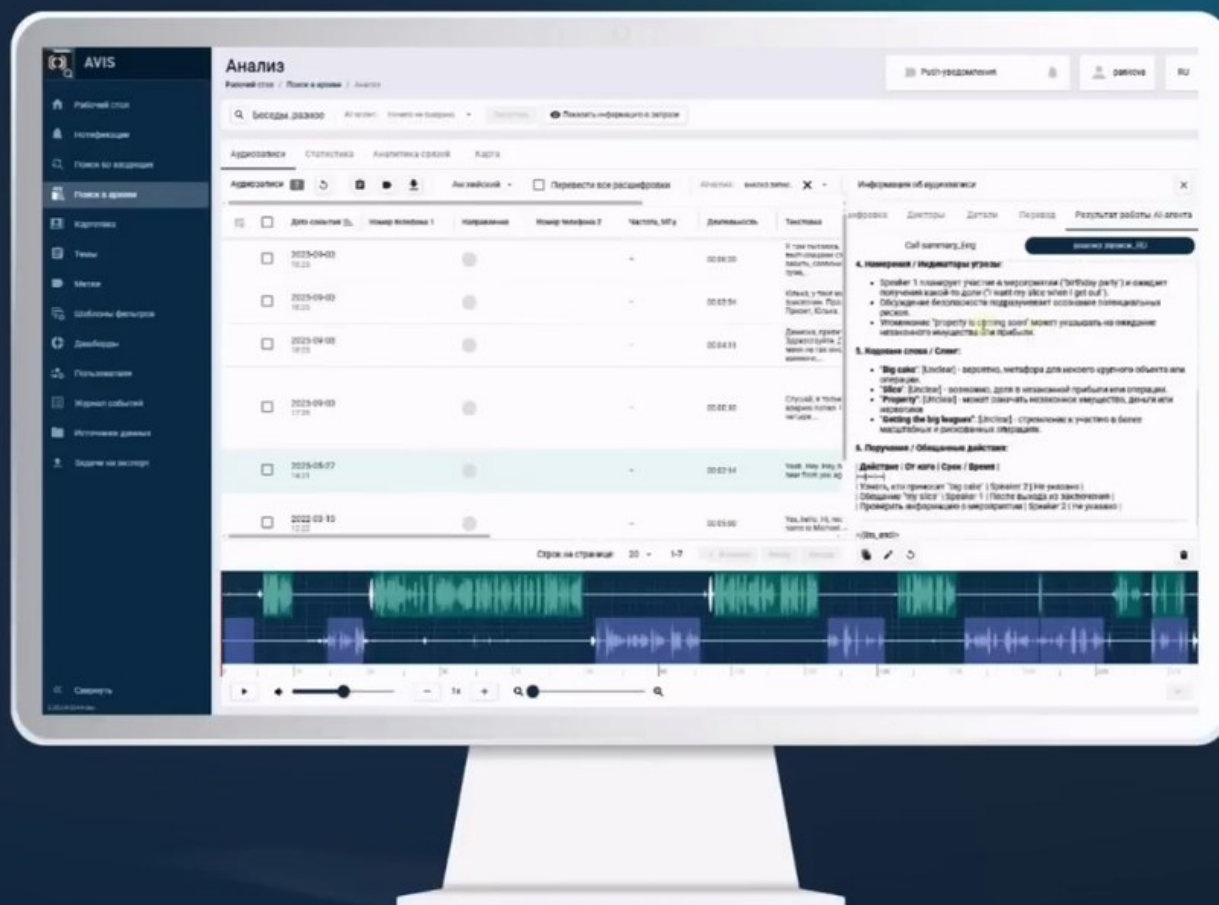
Массовая обработка информации



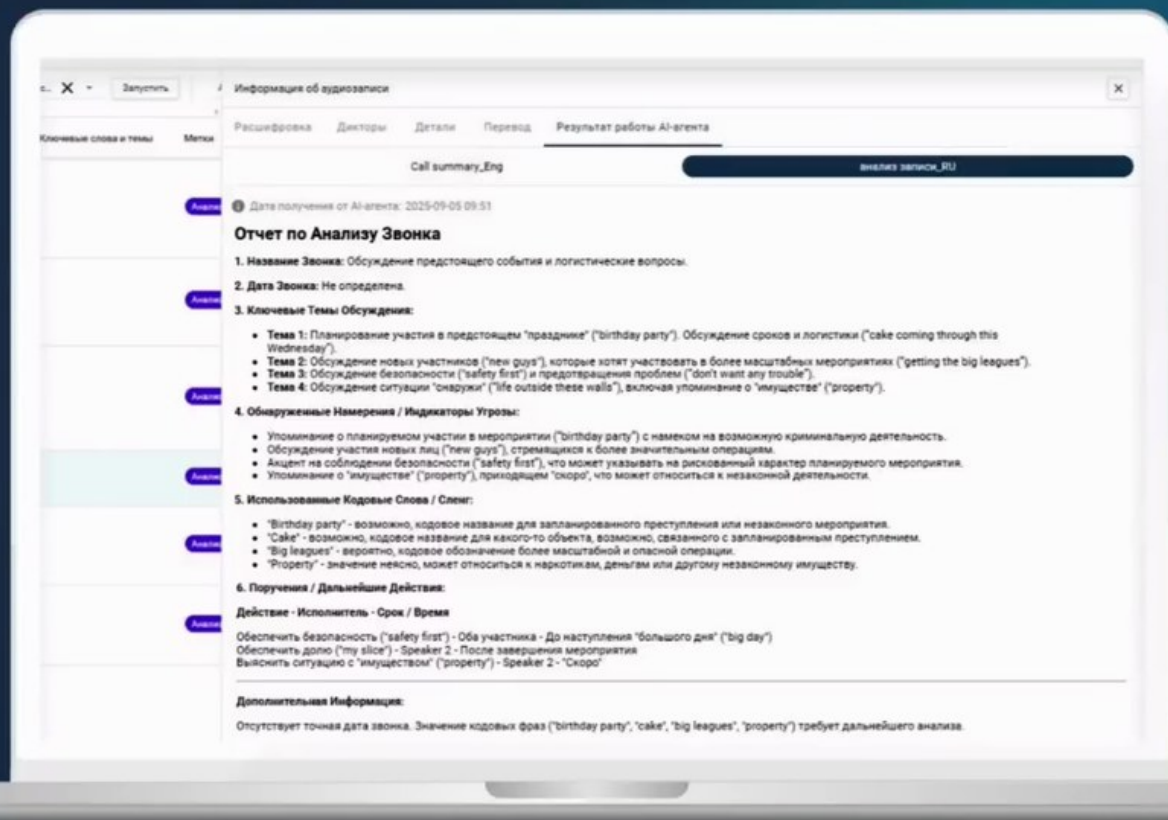
Массовая обработка информации



Массовая обработка информации



Автоаннотирование индивидуальных записей





Выгрузка записей в SIS

Анализ

Рабочий стол / Поиск в архиве / Анализ

🔍 Data KZ + KW Показать информацию о запросе

Аудиозаписи | Статистика | Аналитика связей | Карта

Аудиозаписи 100 🔄 🔍 Поиск в архиве 🏷 Изменить метки 📁 Выгрузка ☐ Перевести все расшифровки Русский - Перетранскрибировать Русский - 🔔 Push-уведомления 👤 admin RU

🔧	➕ Дата события	Номер телефона 1	Направление	Номер телефона 2	Экспорт в ZIP	Имя файла	Источник файла	Группы доступа	Дикторы	Ключевые слова и темы	Метки	Класс
☑	2023-01-24 21:53	10509	➡	10510	Экспорт в ZIP Экспорт в SIS Выгрузить аудио	ссылка16_m10509	GridChannelWebTrans		1 - Cluster_31 2 - Cluster_34	Localization		1 - 0.00s 2 - 0.00s
☑	2023-01-24 21:52	00001	➡	00003		ссылка3_m00001	GridChannelWebTrans		1 - Cluster_1 2 - Cluster_20	Estados policiales... Financia	Control	1 - 0.00s 2 - 0.00s
☑	2023-01-24									Actividad y reuniones...		1 - 0.00s

Предварительный анализ



Извлечение признаков и идентификация

Автоматические
алгоритмы сравнения



Детектирование
речи и шумов



Редактирование
и обработка фонограмм



Идентификационный
визард



Экспертные
методы сравнения



Расчет
характеристик сигнала



Визуализация
и анализ фонограмм



Модуль EdiTracker



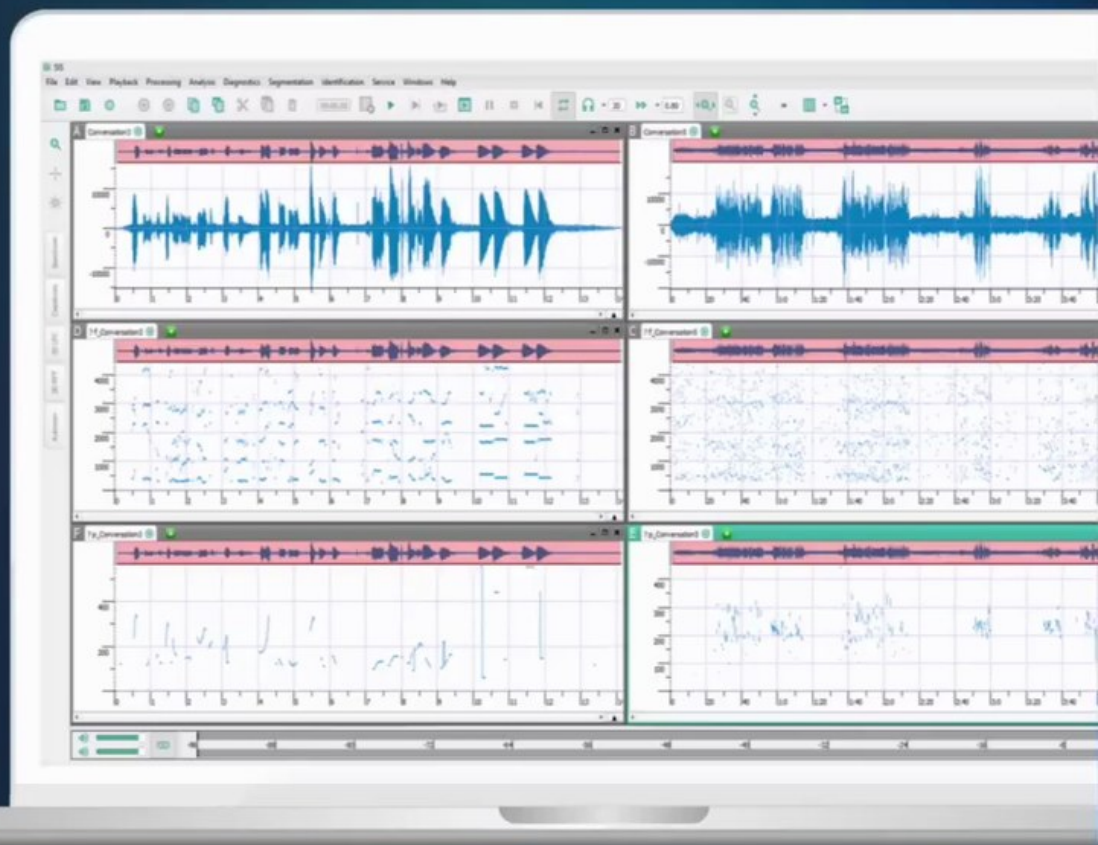
Диагностика
аутентичности



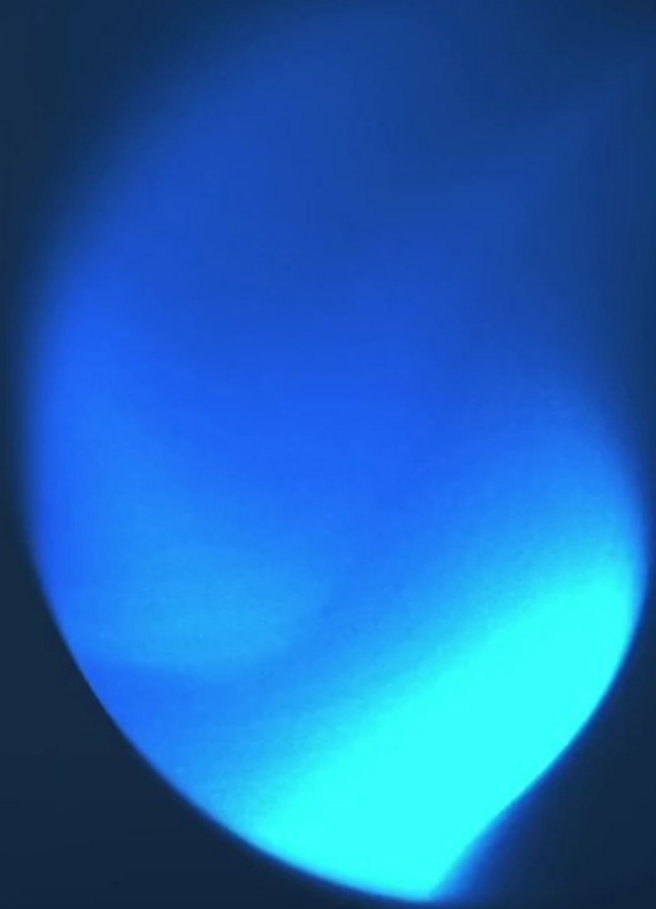
Заключение
эксперта



Модуль
«Супер лупа»



Пример 2



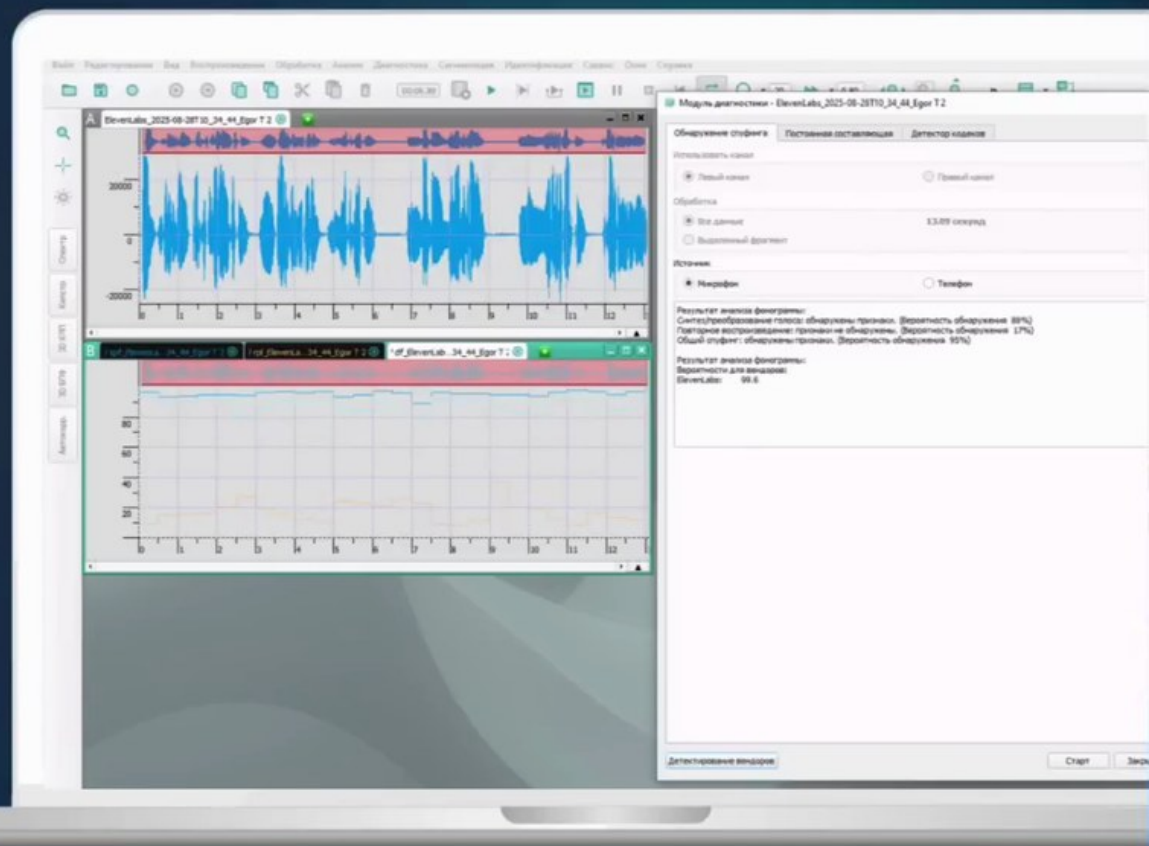
Детектор спуфинга

Модуль детектирования* оценивает вероятность спуфинга (синтез/конверсия, реплей, общая оценка) на фонограмме.

Рассчитывается общая оценка и отдельные оценки по типам атак.

Формируются три графика спуфограммы с отображением значений скользящих оценок в соответствии со временем сигнала.

При наличии оценки вероятности синтеза от 50% можно включить дополнительную функцию определения вендора синтеза.



* Поставляется под отдельной лицензией

Проверка спуфинга или проверка аутентичности

Проверка на спуфинг — идентификация и выявление поддельных медиафайлов с помощью алгоритмов анализа нестыковок и технологий детекции дипфейков.

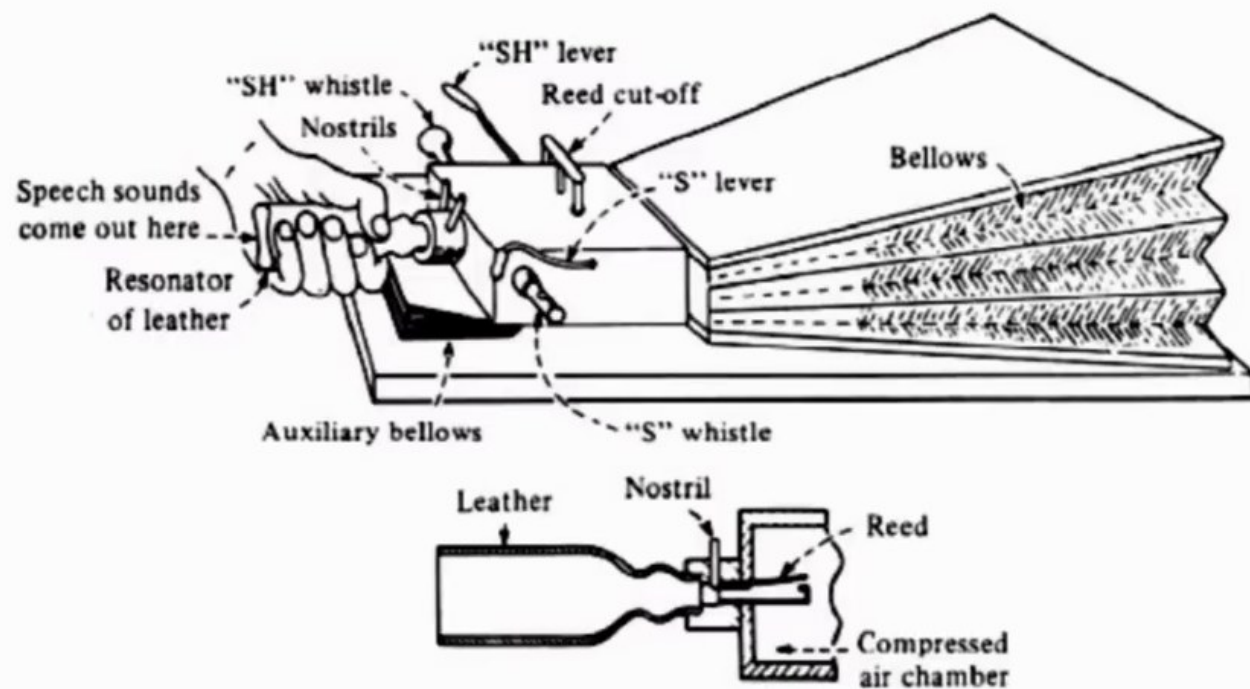
Проверка аутентичности — подтверждение оригинальности и целостности цифрового контента:

- анализ метаданных;
- проверка цифровых подписей и водяных знаков;
- сравнение с достоверными источниками;
- обнаружение дипфейков.

Обнаружение дипфейков направлено на выявление поддельных медиафайлов, в то время как проверка подлинности подтверждает оригинальность и целостность цифрового контента.



История спуфинга



Mark J. Jones, Rachael-Anne Knight (Eds.) The Bloomsbury Companion to Phonetics. 2013

Некоторые примеры

1920

1979

...

2000



2023

2024



Fidelity Voice Chess Challenger (1979)
первые говорящие шахматы

https://www.ismenio.com/chess_fidelity_vcc.html

Простой генератор речи

<https://ieeexplore.ieee.org/document/861820>

Простое ИИ-клонирование голоса

Google search: voice clone

Продвинутое системы клонирования голоса с эмоциями

<https://www.cnn.com/2024/01/22/politics/fake-joe-biden-robocall/index.html>

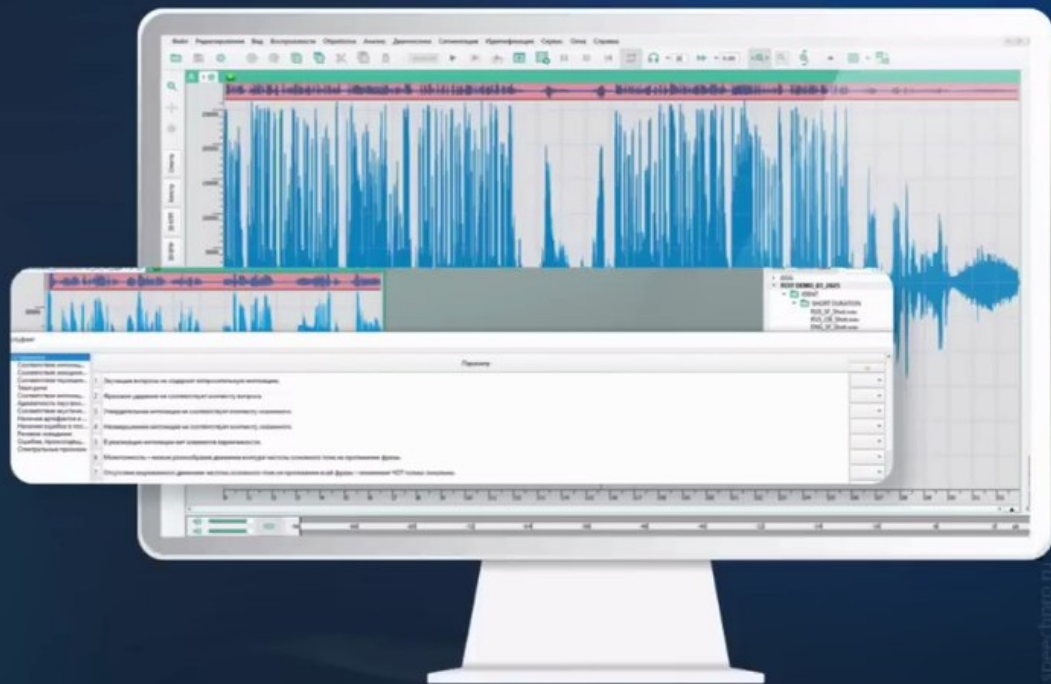


Возможности SIS

Экспертный метод определения спуфинга

Оценка спуфинга

- ➔ Автоматический детектор формирует первичную оценку записи
- ➔ Эксперт фиксирует признаки спуфинга в интерактивной таблице артефактов*
- ➔ Итог — отчёт, сохраняемый в проект



* Поставляется под отдельной лицензией

Возможности SIS

Модуль EdiTracker

EdiTracker позволяет выявить признаки нарушения аутентичности с помощью следующих методов:

- ➔ Исследование технических характеристик записывающего устройства
- ➔ Обнаружение следов цифровой предобработки сигнала
- ➔ Обнаружение стационарных гармоник и сканирование их фазы
- ➔ Сканирование фонового шума
- ➔ Аудитивный анализ

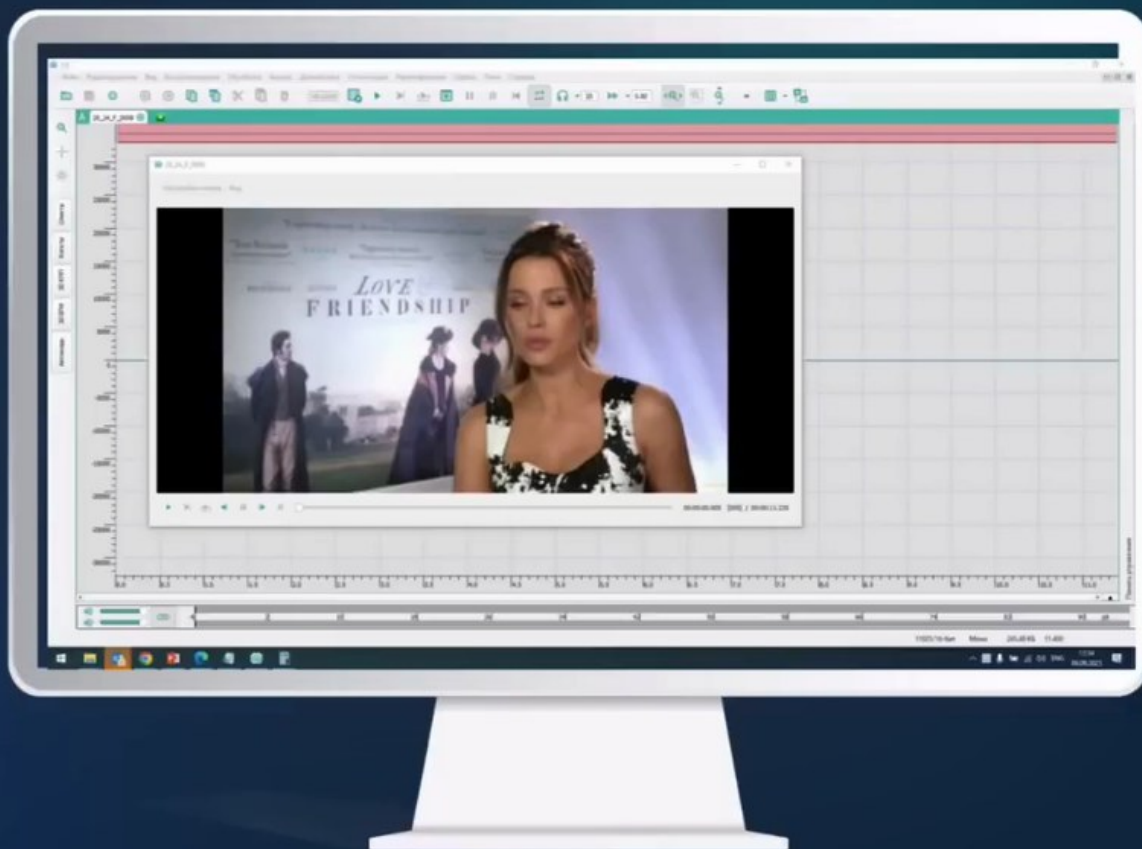


Пример 3



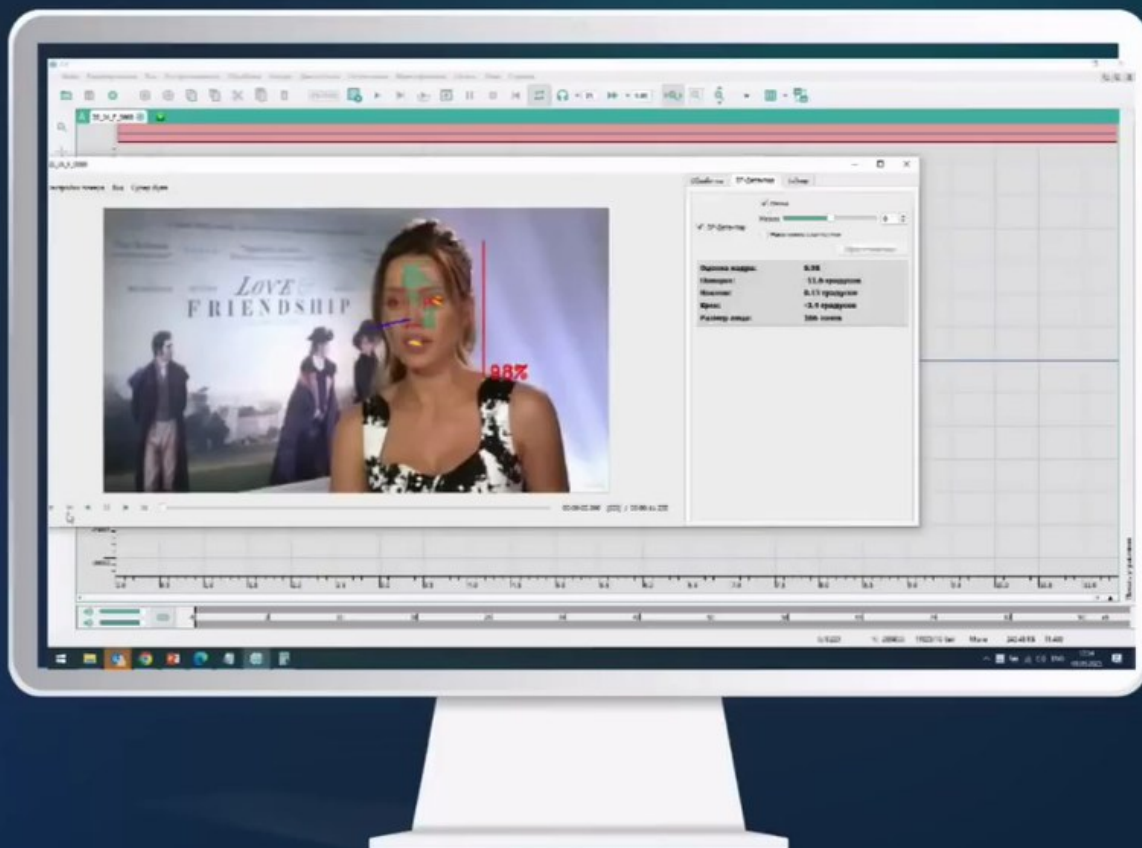
Анализ дипфейков

- ➔ Проводится анализ видеокадров медиафайла
- ➔ Определена высокая вероятность дипфейка для синтетического видео
- ➔ Покадровый анализ показывает оценку для каждого кадра
- ➔ После анализа выводится усредненная высокая вероятность наличия дипфейка



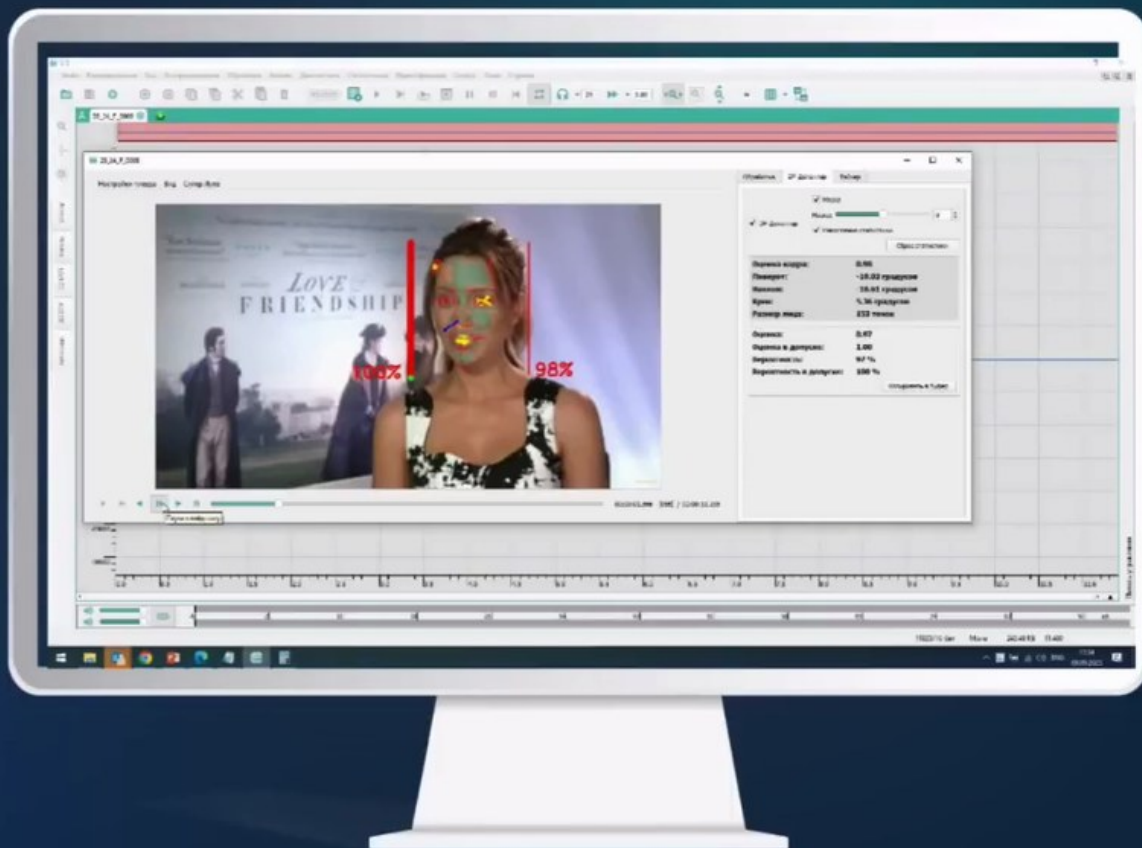
Анализ дипфейков

- Проводится анализ видеок кадров медиафайла
- Определена высокая вероятность дипфейка для синтетического видео
- Покадровый анализ показывает оценку для каждого кадра
- После анализа выводится усредненная высокая вероятность наличия дипфейка



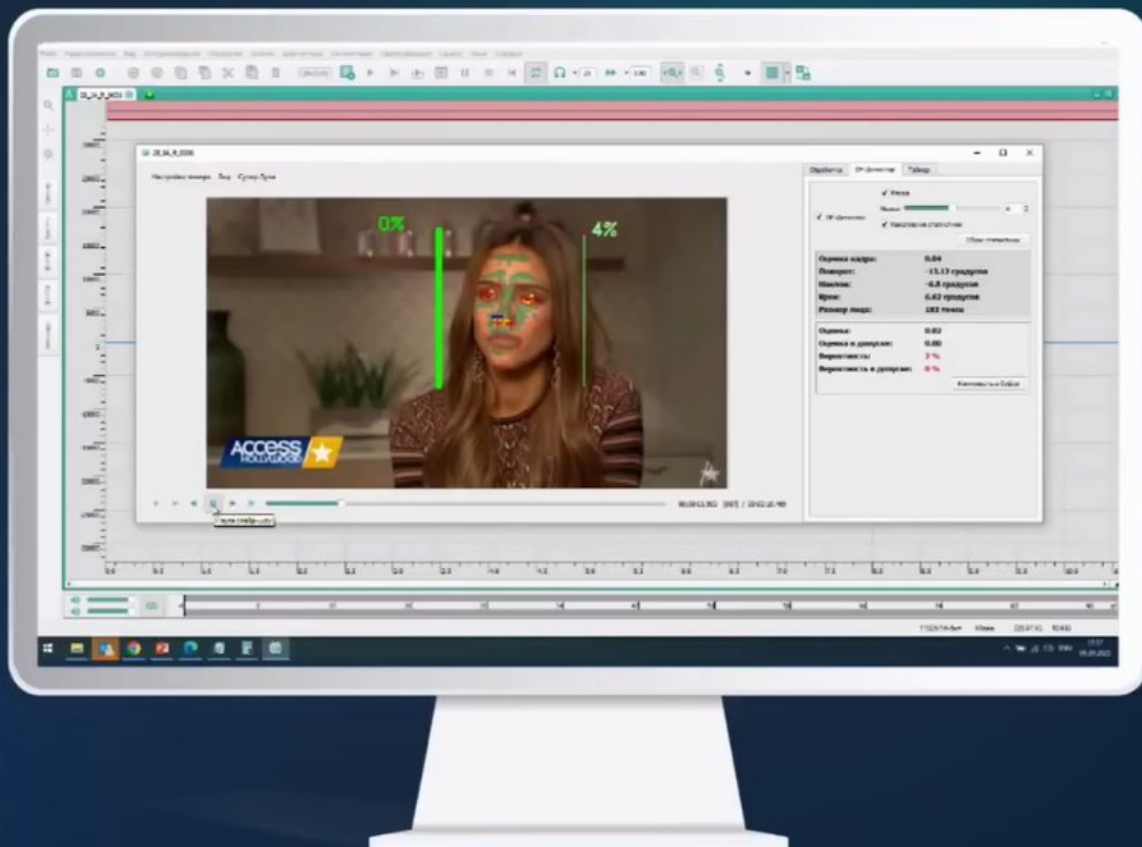
Анализ дипфейков

- Проводится анализ видеок кадров медиафайла
- Определена высокая вероятность дипфейка для синтетического видео
- Покадровый анализ показывает оценку для каждого кадра
- После анализа выводится усредненная высокая вероятность наличия дипфейка



Анализ дипфейков

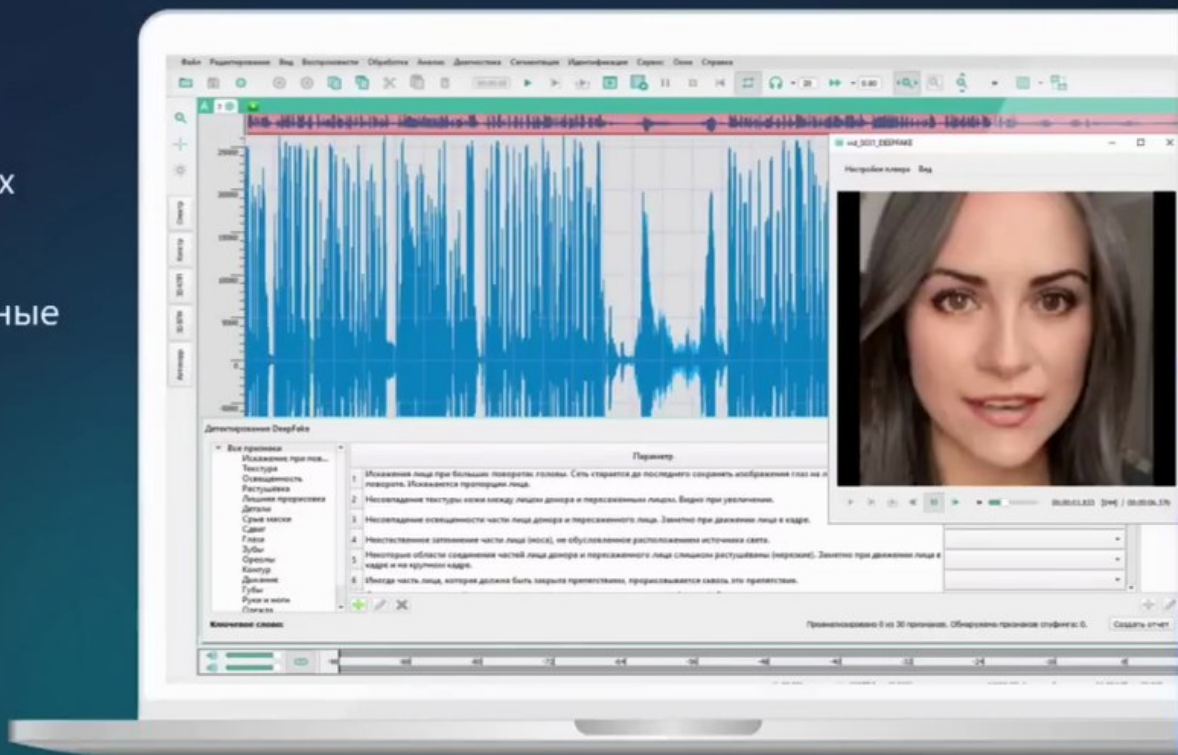
- Проводится анализ видеок кадров медиафайла
- Определена низкая вероятность дипфейка для оригинального видео
- Покадровый анализ показывает оценку для каждого кадра
- После анализа выводится усредненная низкая вероятность наличия дипфейка



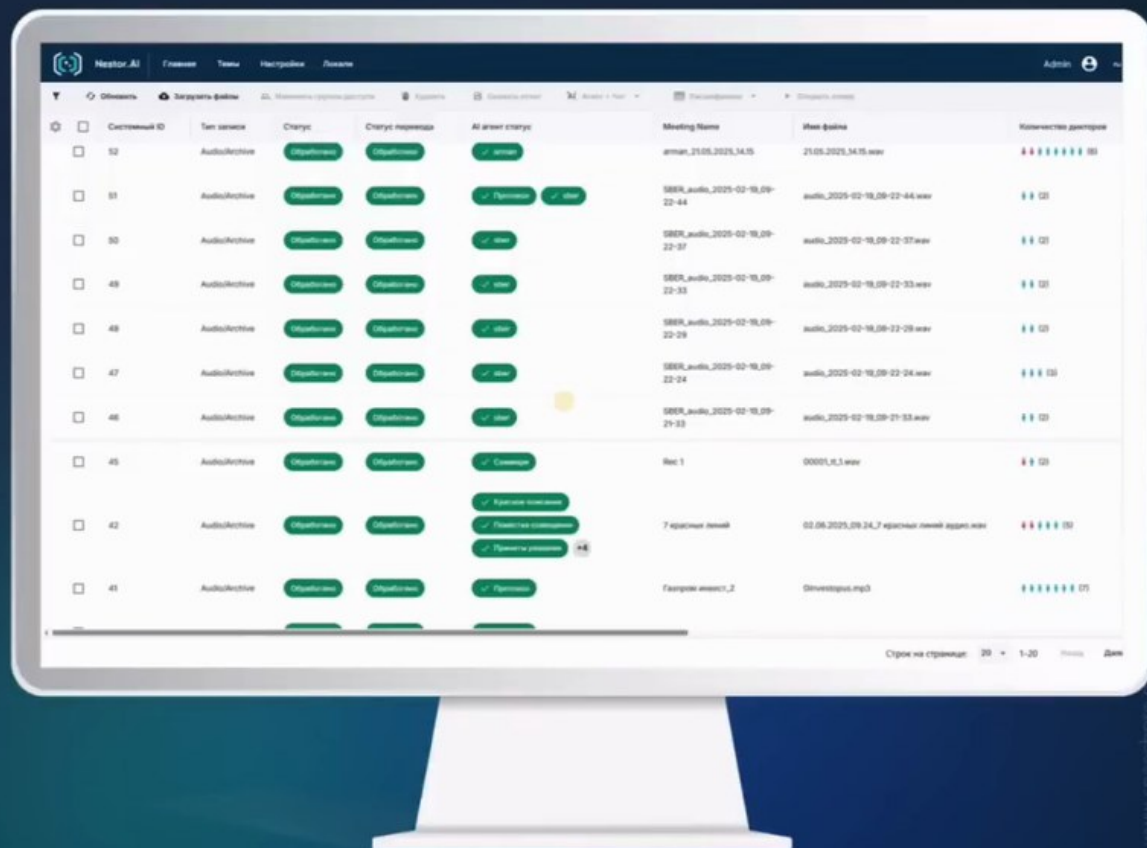
Возможности SIS

Экспертный метод определения дипфейков

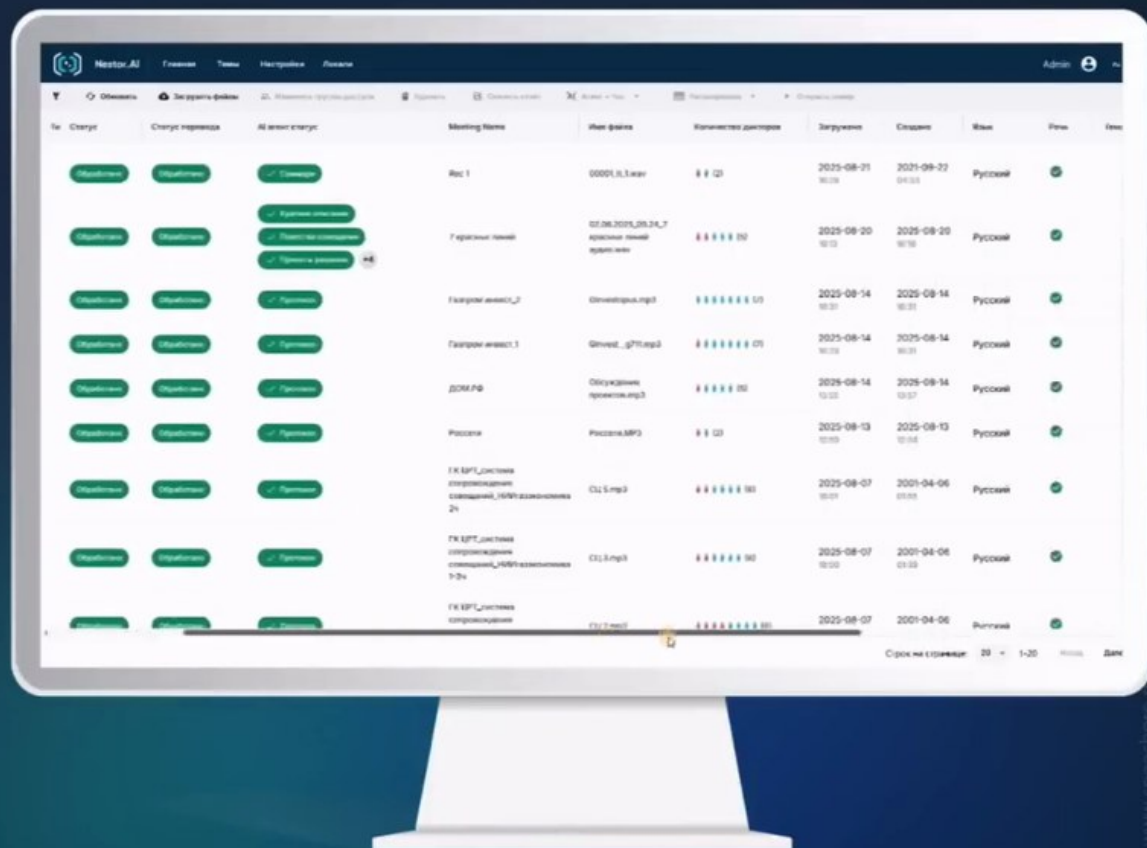
- ➔ Автоматическая оценка может быть дополнена **экспертным анализом признаков дипфейка*** на видеокадрах
- ➔ Эксперт использует интерактивную таблицу артефактов, отмечая характерные искажения не только на лице, но и по всему кадру
- ➔ Результатом становится отчёт, который сохраняется в проекте



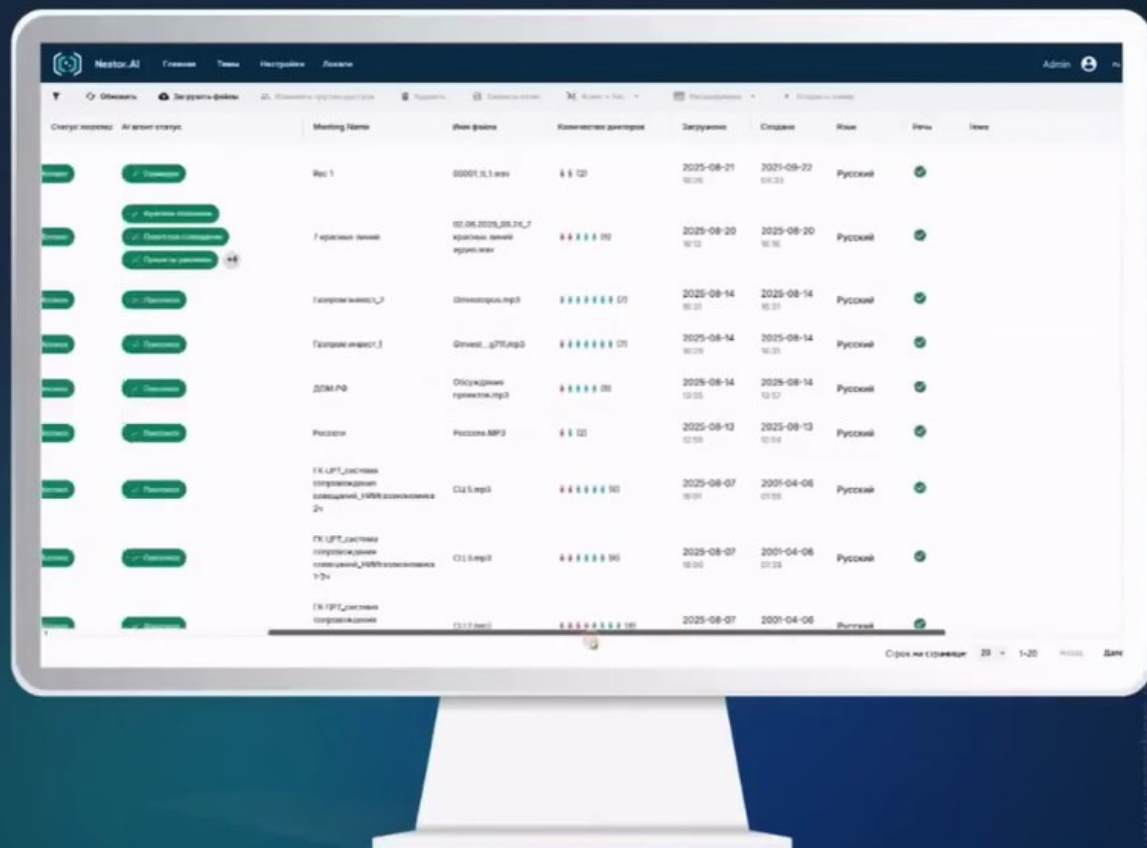
При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой



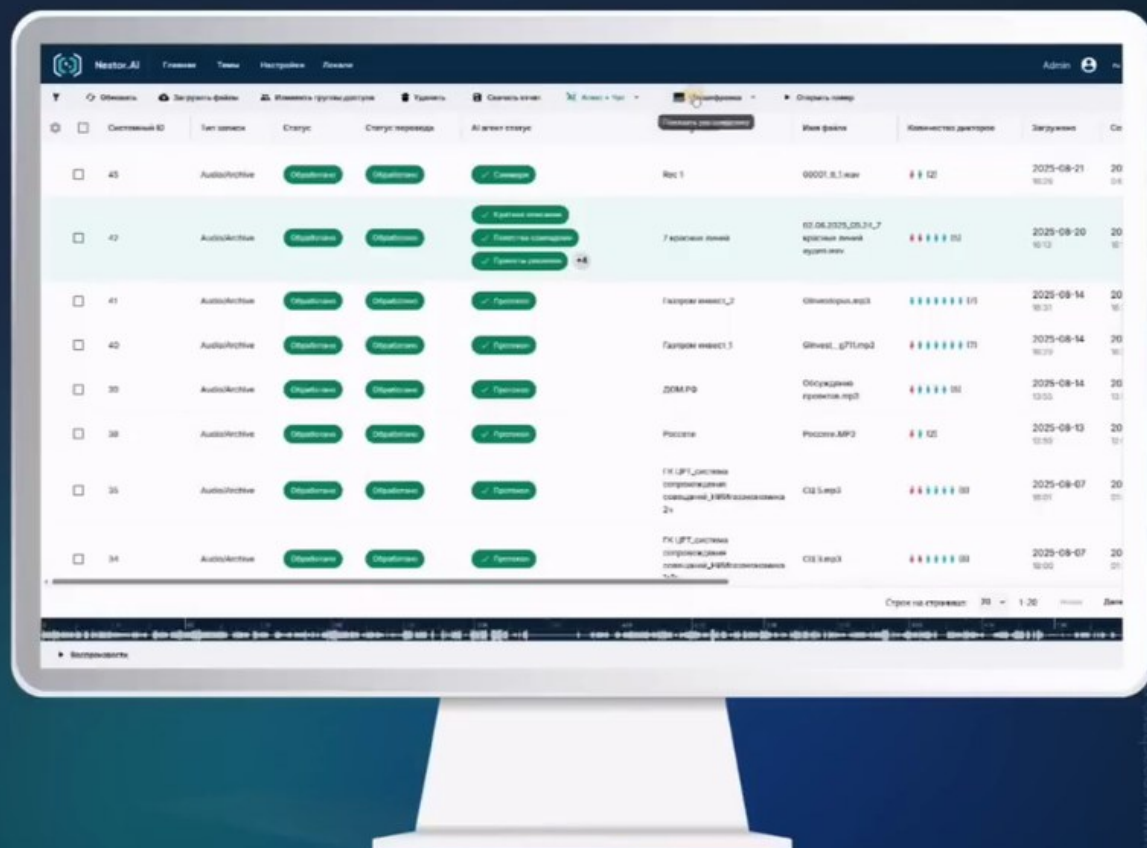
При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой



При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой

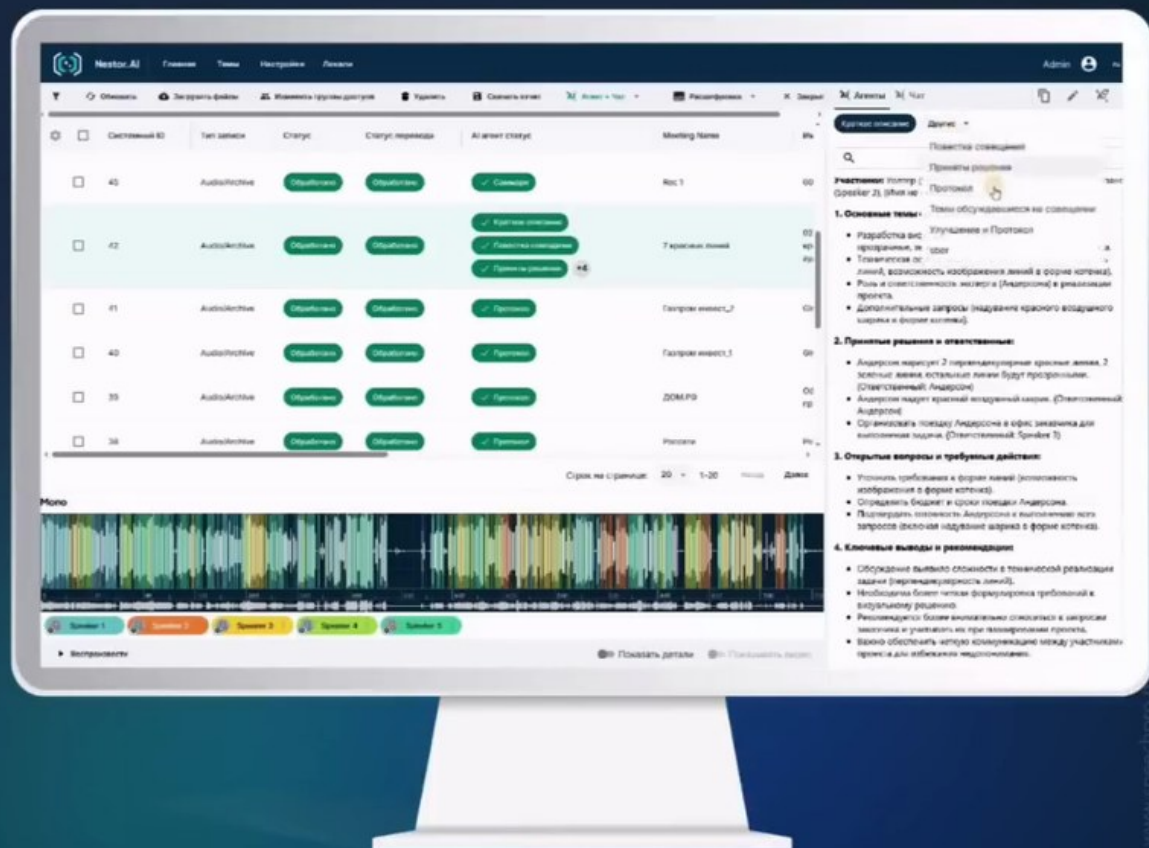


При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой

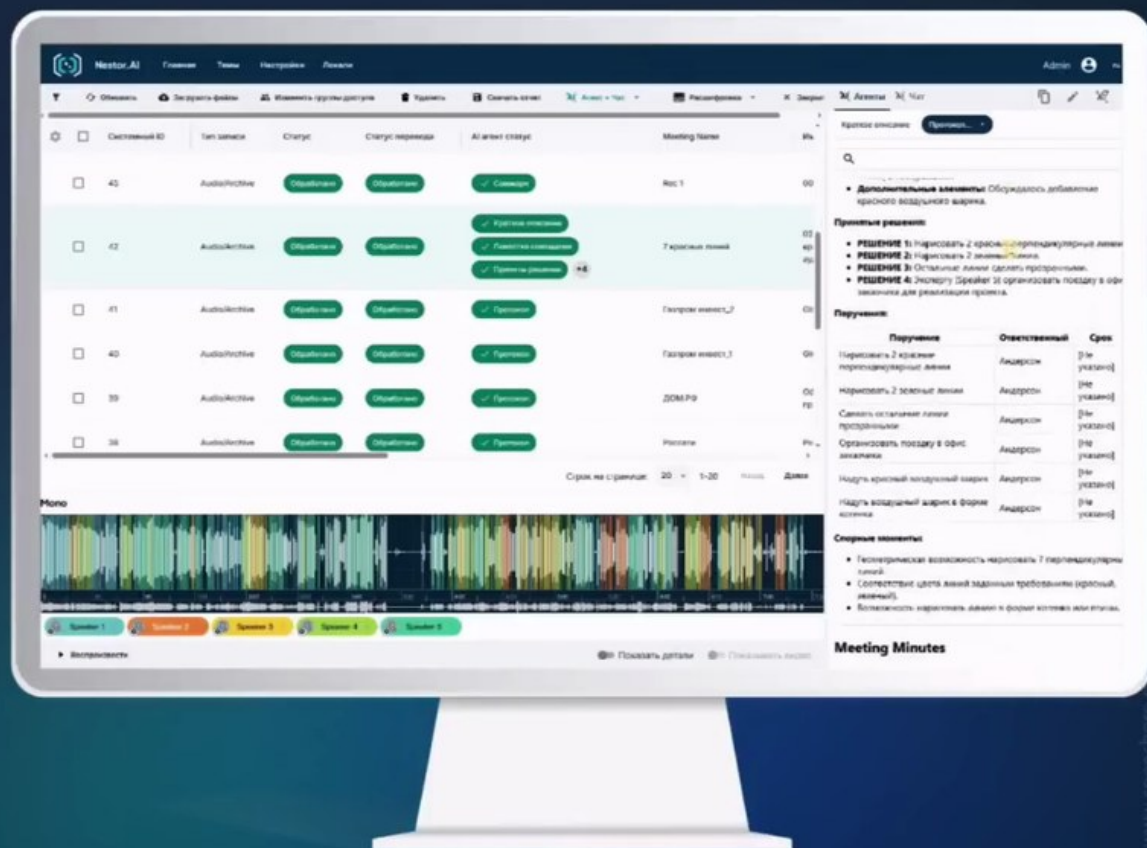


The screenshot displays the Nestor AI web application. At the top, there's a navigation bar with the logo and menu items: Главная, Темы, Настройка, Личные. The main header includes a search bar and several tabs: Обзор, Задачи, Задачи-фильтры, Задачи-список, Задачи, Статистика, and Задачи-комментарии. Below the header is a table with columns: ID, Тип задачи, Статус, Статус-комментарий, AI-анализ, Имя файла, Количество действий, and Дата. The table lists several tasks, each with a checkbox, ID, type (Audio/Text), status buttons (Обзор, Комментарий, Проверка), AI analysis status (Проверка, Проверка-комментарий, Проверка-результат), file name, number of actions, and date. Below the table, there's a section labeled 'Мониторинг' showing a timeline with colored bars representing different tasks or activities over time. At the bottom, there's a footer with a search bar and a button 'Поискать детали'.

При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой



При проведении судебных разбирательств по аудиоподделке использовалась система Nestor.AI, которая автоматически превратила речь заседания в текст, а ИИ помощник подготовил краткую сводку содержания и сохранил ее в документ строгой формы вместе со стенограммой



Заключение

Продукты группы компаний ЦРТ работают на разных этапах криминалистического процесса, позволяют использовать речевую биометрию, лицевую биометрию и интеллектуальные речевые технологии для решения задач поиска целевых материалов, а также получения криминалистических заключений